

DIE NUTZUNG VON EVALUATIVEN INFORMATIONEN FÜR SCHUL- UND UNTERRICHTSENTWICKLUNG

Herbert Altrichter
Linz School of Education

ÜBERSICHT

1. 'Modernisierung' des Schulwesens und die Logik der 'evidenzbasierten Steuerung'
2. Forschung zur 'evidenzbasierten Steuerung'
 1. Schulinspektion
 2. Bildungsstandards und Datenfeedback
3. Zusammenfassung und Ausblick: Gehen QIBB und SQA darüber hinaus?

PHASEN DER 'SCHULMODERNISIERUNG'

- *Phase 0: ‚Doppelsteuerung‘ (Brüsemeister, 2004)*
 - Staatliche input-Regulierung, bürokratisch-hierarchische Administration, generelle Regeln, gleiche Ressourcen
 - Beträchtliche individuelle und gruppenbezogene Autonomie für die Lehrpersonen bei der Umsetzung der Regeln
 - 1980er Jahre: „bildungspolitisch eigenartige unbewegte Zeit“ (Fend 2006)
- *Phase I: Öffnung von Spielräumen – 1. Hälfte 1990er Jahre*
 - Autonomie/Gestaltungsspielräume/Schulentwicklung, schulautonomes Curriculum, innere Organisation
 - Hoffnung auf Entwicklung durch professionelle Kompetenzen und Energie der LehrerInnen

PHASEN DER 'SCHULMODERNISIERUNG'

- *Phase 2: Innerschulisches Management - 2. Hälfte der 1990er Jahre*
 - *Interne Steuerungsinstrumente: Stärkung des Managements, Schulprogramm, Selbstevaluation und Qualitätsmanagement*
→ Konstituierung der Schule als „Betrieb“
 - *Erste externe Steuerungsinstrumente: Vergleichsarbeiten, Aufgabenbeispiele* → „Vielfalt orchestrieren“
- *Phase 3: PISA-Schock und schulübergreifende Steuerung 2001 –*
 - *Verstärkung externer Regulierung*
 - *einheitliche Bildungsstandards, vergleichende Leistungsüberprüfung und Datenfeedback; Qualitätsrahmen + neue Schulinspektionen; Evaluationsagenturen, Nationale Bildungsberichte*

LEITIDEEN EINES „EVIDENZBASIERTEN STEUERUNGSMODELLS“

- *Ziele*: formuliert *Erwartungen für die Leistungen* des Schulsystems und kommuniziert sie entschiedener als zuvor
 - „Bildungsstandards“
 - „Qualitätsrahmen“ von Schulinspektionen
- *Ist-Soll-Vergleich*: soll „*Evidenz*“ bzgl. Leistungserwartungen produzieren
 - standardbezogenen Leistungstests (Lernstandserhebungen)
 - „neue Schulinspektionen“
 - Errichtung von Instituten der Bildungsforschung und Qualitätssicherung.
- *Entwicklung*: Diese „*Evidenz*“ soll *Systementwicklung und -verbesserung* auf allen Systemebenen stimulieren und orientieren (d.h. auf die richtige Ziele richten).
 - Datenfeedback: Rückmeldung von Lernstandserhebung und Inspektionsberichten
 - Regionale und Nationale Bildungsberichte
- *Transparenz/Öffentlichkeit*: Verschiedene stakeholder der Schulen sollen einbezogen werden, um externe Aufmerksamkeit (Druck/Unterstützung) zu erhöhen
 - Veröffentlichung von Bildungsstandardergebnissen, Berichte an Eltern
- beansprucht *Orientierungskraft für alle Ebenen* des „Mehrebenensystem Schule“
 - Rückmeldung an Lehrperson, Schülerinnen/Eltern, Einzelschule, Aufsichtsregion, Ministerium

PHASEN DER ,SCHULMODERNISIERUNG‘

- *Phase 4: Ernüchterung, Einschränkungen und Kritik - ???*
 - Z.B: Schulinspektion: Niedersachsen (2013): dialogischer Ansatz, weniger Qualitätsbereiche, stärkere Berücksichtigung von Prozessen
 - Schleswig-Holstein 2010, Rheinland-Pfalz (2016): Abschaffung der Schulinspektion, Zielvereinbarungen
 - NL (2008): proportionate, risk-based inspection
- Europäische Tendenzen:
 - Stärkere Kostenüberlegungen
 - Fokussierung der Evaluationstätigkeiten auf ‘schwache Schulen’
 - mehr Spielräume für ,professionelle Schulen‘

IM KERN: DAS „EVALUATIVE ARGUMENT“

- = dass (politische und praktische) Entscheidungen über Weiterführung und Weiterentwicklung eines Programms
- besser (begründeter und rationaler) getroffen werden können,
- wenn die bisherige Praxis dieses Programms und seine Ergebnisse
- vor einem Satz explizit gemachter Kriterien
- nach geregelter, durch die (Sozial-)Wissenschaften legitimierten Methoden
- untersucht und bewertet werden.

ÜBERSICHT

1. 'Modernisierung' des Schulwesens und die Logik der 'evidenzbasierten Steuerung'
2. Forschung zur 'evidenzbasierten Steuerung'
 1. Schulinspektion
 2. Bildungsstandards und Datenfeedback
3. Zusammenfassung und Ausblick: Gehen QIBB und SQA darüber hinaus?

WELCHE *WIRKUNGEN SOLLEN* SCHULINSPEKTIONEN IN VERSCHIEDENEN EUROPÄISCHEN LÄNDERN AUF SCHULEN AUSÜBEN? AUF WELCHE WEISE (DURCH WELCHE PROZESSE UND KONTEXTFAKTOREN) SOLLEN DIESE WIRKUNGEN ERREICHT WERDEN?

Analyse von Programmtheorien (Leeuw 2003)

- Rekonstruktion der Annahmen über kausale Mechanismen und beabsichtigte Effekte
- *Dokumentenanalyse* von Inspektionsregelungen und einschlägigen rechtlichen Dokumenten
- *Interviews* mit Inspektor/innen

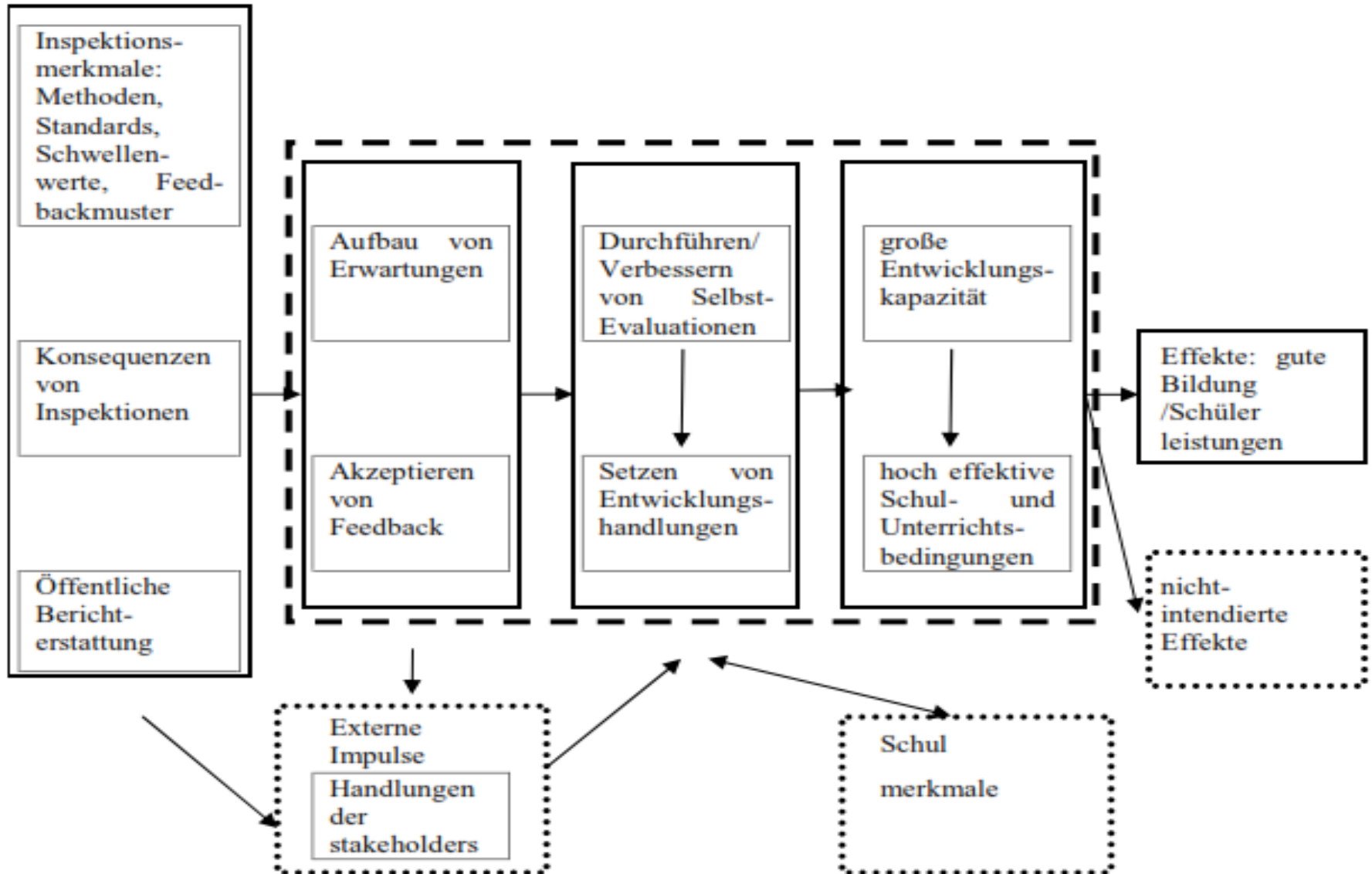
- eine ‘Programmtheorie’ für jedes Land
- Cross-case Analyse: “*konzeptuelles Modell*”, das den “konzeptuellen Kern” von Inspektion repräsentiert

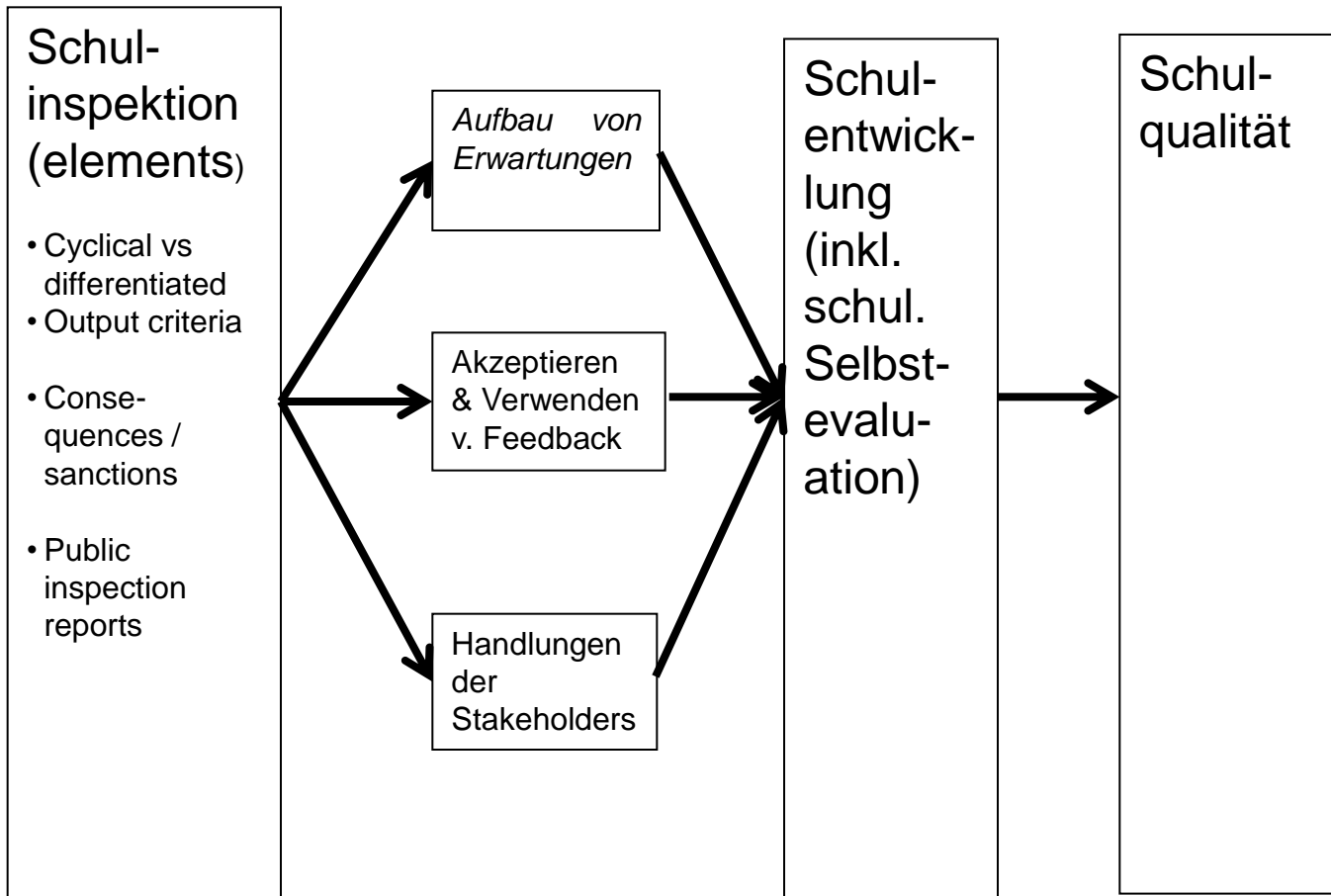
ENTWICKLUNG EINES WIRKUNGSMODELLS VON SCHULINSPEKTIONEN DURCH VERGLEICHENDE ANALYSE DER PROGRAMMTHEORIEN

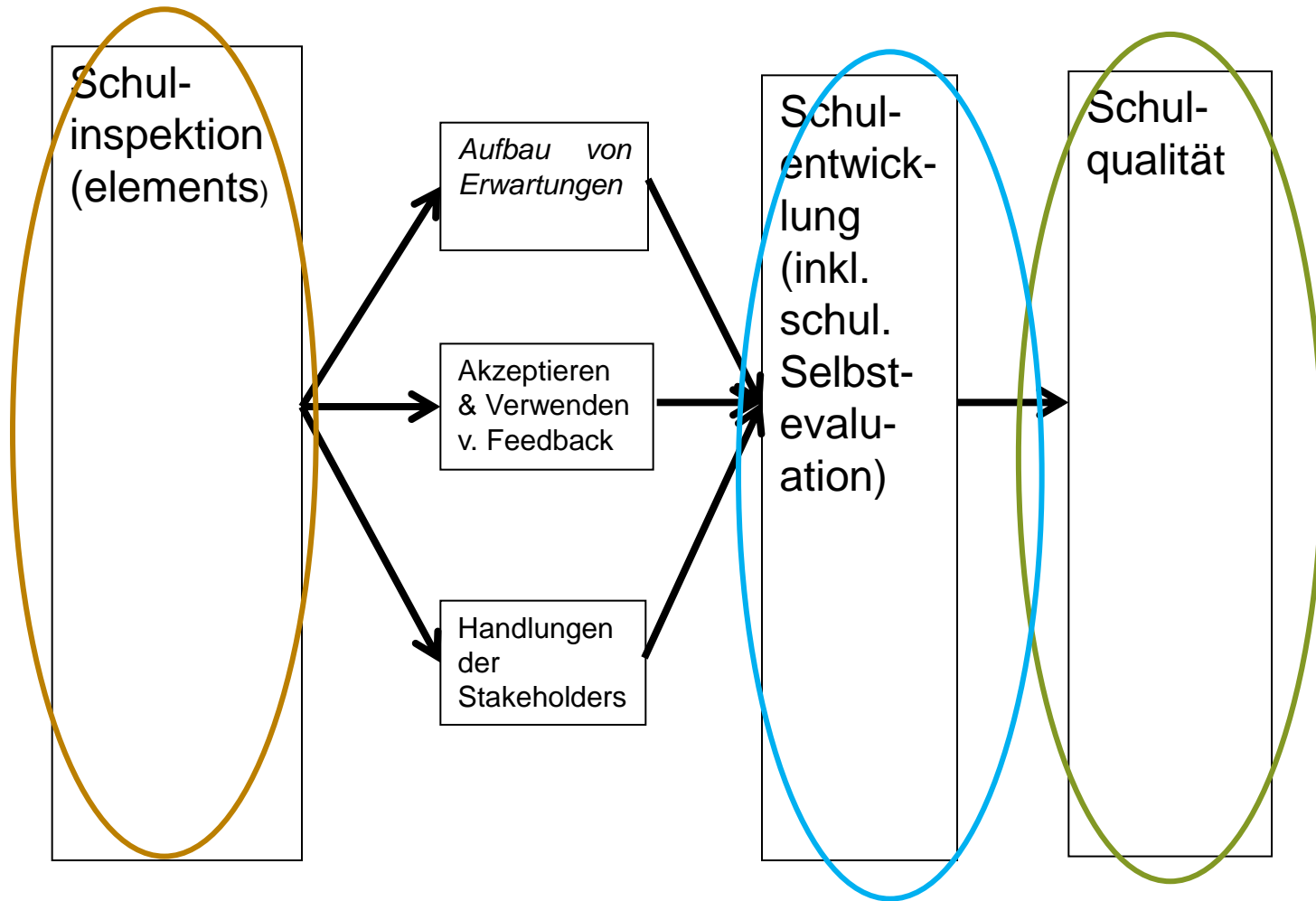
- Wirkungsmodell aus dem Vergleich der nationalen „Programmtheorie“ Einzelmodelle
- repräsentiert offiziellen Erwartungen und Wirkungsannahmen -- ein normatives Modell

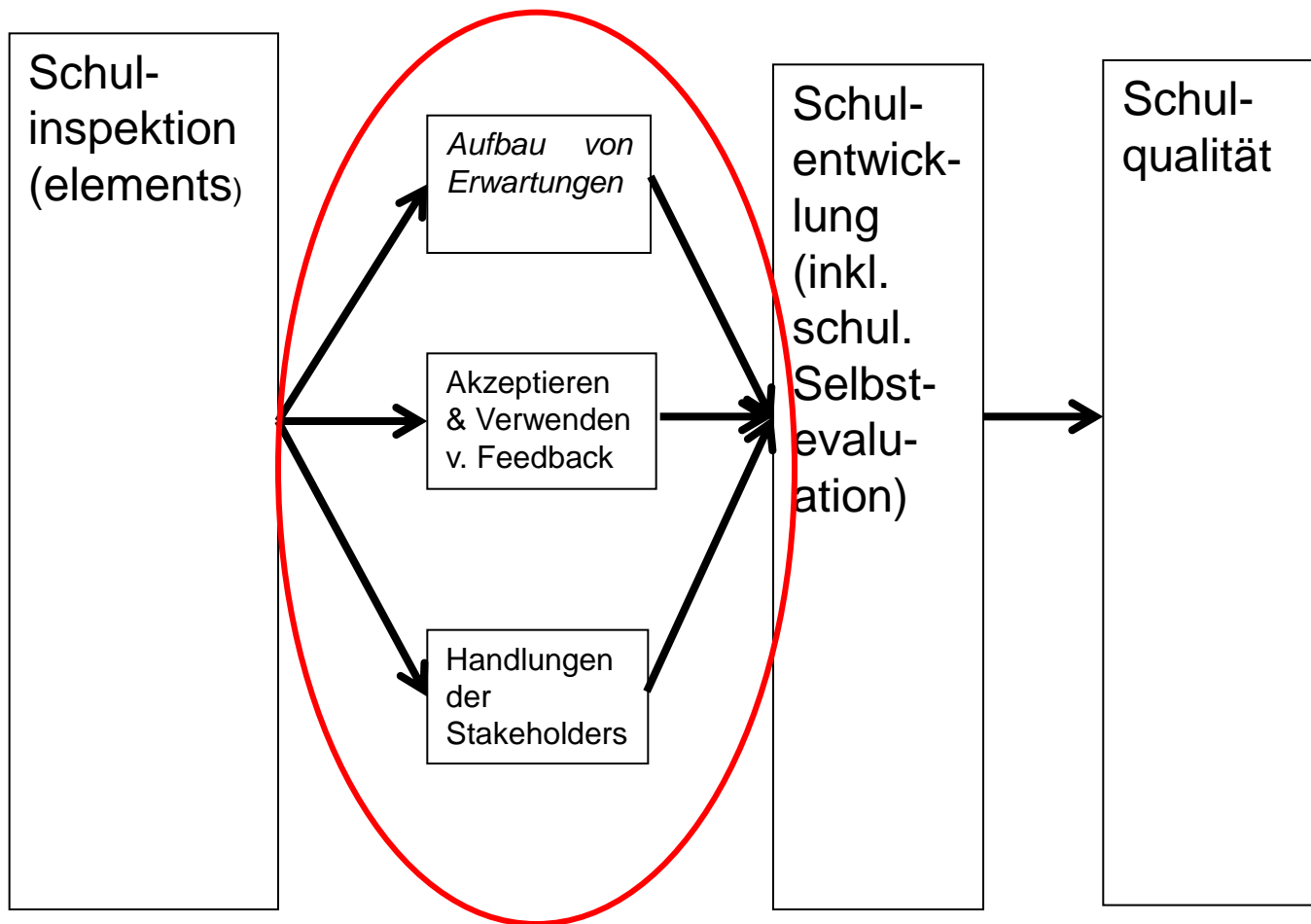
ABB. 3: WIRKUNGSMODELL VON SCHULINSPEKTIONEN

(EHREN, ALTRICHTER, MCNAMARA & O'HARA 2013)









„VERMITTELNDE MECHANISMEN“ ZUR ERZIELUNG DER INTENDIERTEN EFFEKTE

- ***Aufbau von Erwartungen:*** Inspektionen kommunizieren Erwartungen über schulische Ziele und Prozesse durch Standards, Auftreten in Besuchen, Berichte und Konsequenzen
- ***Akzeptieren und Verarbeiten von Feedback:*** Inspektionen „wirken“ durch ihre Feedbackberichte, wenn Schulen diese Botschaften akzeptieren und als Leitlinie für künftige Handlungen nehmen
- ***Aufmerksamkeit und Handlungen der stakeholders:*** Inspektionen machen die stakeholders der Schule auf Qualitätsmerkmale und Stärken und Schwächen aufmerksam und motivieren sie, durch Unterstützung oder Druck auf Schulentwicklung zu dringen.

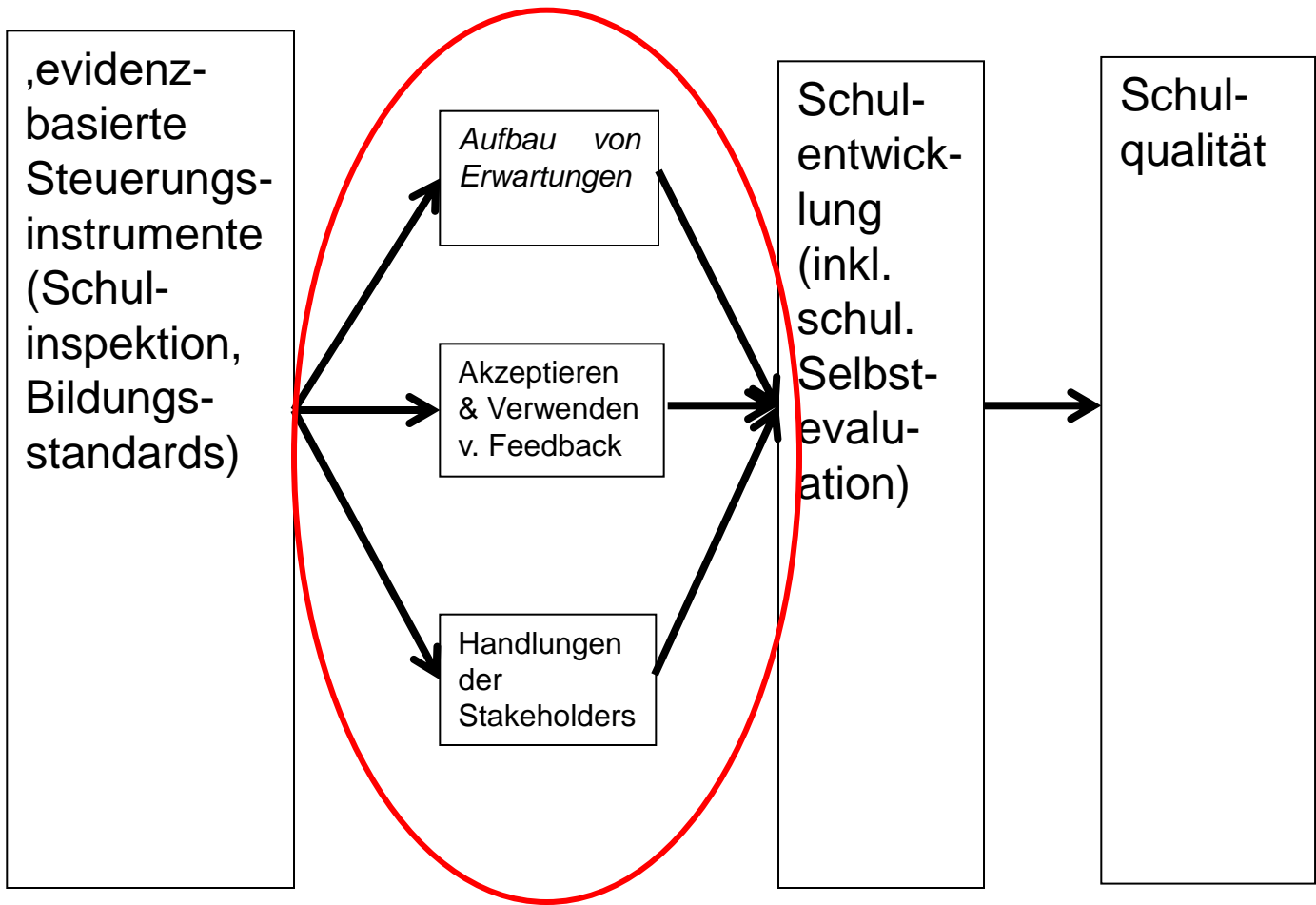
LEITIDEEN EINES „EVIDENZBASIERTEN STEUERUNGSMODELLS“

Aufbau von
Erwartungen

- *Ziele*: formuliert *Erwartungen für die Leistungen* des Schulsystems und kommuniziert sie entschiedener als zuvor
- *Ist-Soll-Vergleich*: soll „*Evidenz*“ bzgl. Leistungserwartungen produzieren
- *Entwicklung*: Diese „*Evidenz*“ soll *Systementwicklung und -verbesserung* auf allen Systemebenen stimulieren und orientieren (d.h. auf die richtige Ziele richten).
- *Transparenz/Öffentlichkeit*: Verschiedene stakeholder der Schulen sollen einbezogen werden, um externe Aufmerksamkeit (Druck/Unterstützung) zu erhöhen
- beansprucht *Orientierungskraft für alle Ebenen* des „Mehrebenensystem Schule“

Akzeptieren &
Verwenden v.
Feedback

Handlungen der
Stakeholders



EU-LLP PROJEKT „IMPACT OF SCHOOL INSPECTIONS ON TEACHING AND LEARNING“ (ISI-TL)

Forschungsteams aus acht europäischen Ländern: NL, UK/ENG, IRL, SWE, CZE, AUT (Steiermark)

Online-Befragung von 2300 Primar- und Sekundar-Schulleiter/innen im Längsschnitt

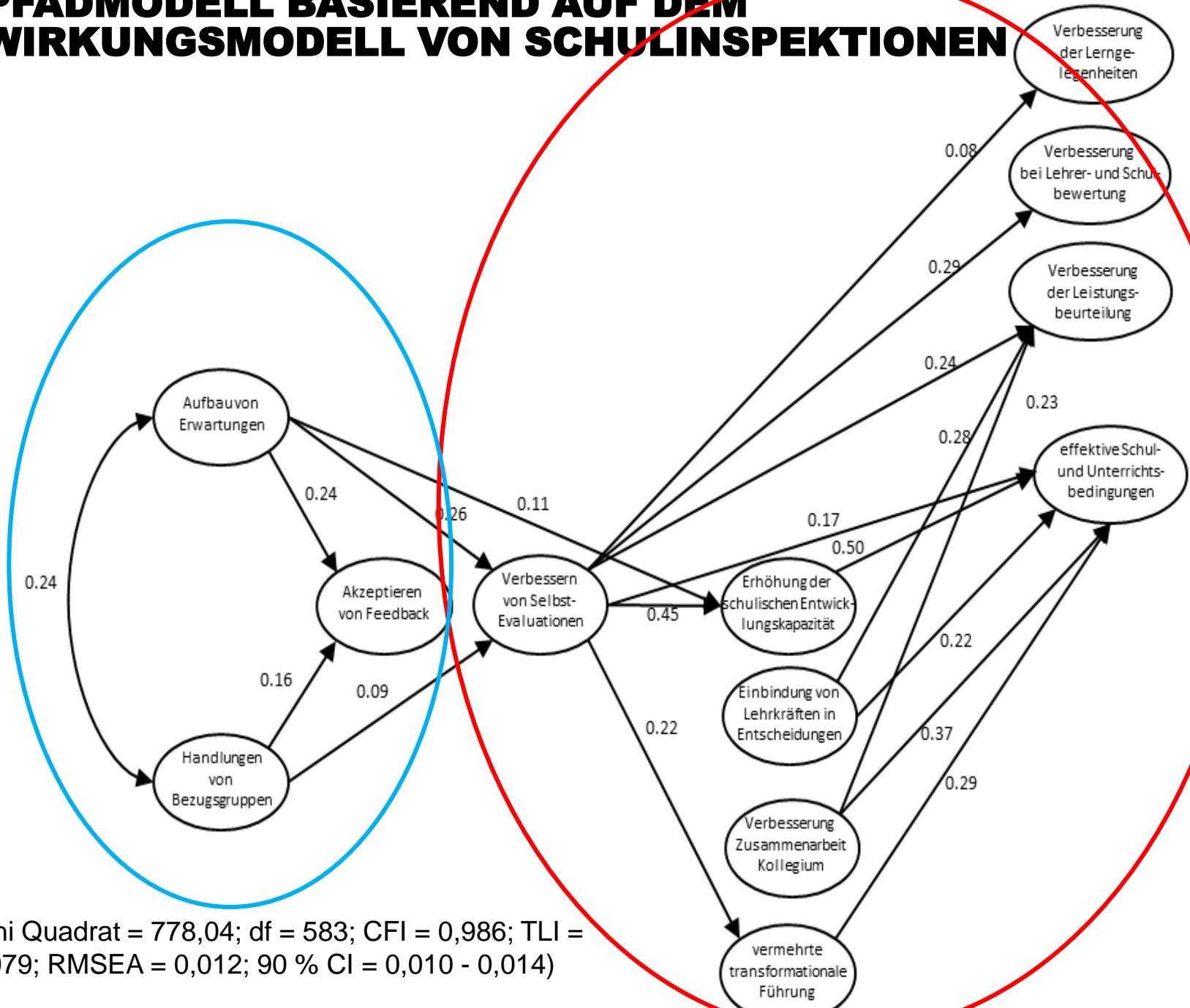
- ISI-TL1 November 2011 – Jänner 2012
- ISI-TL2 November 2012 – Jänner 2013
- ISI-TL3 November 2013 – Jänner 2014

Durch welche Mechanismen fördern Schulinspektionen Entwicklungsaktivitäten (bzw. unerwünschte Konsequenzen)?

Gustafsson, J.-E., Ehren, M.C.M., Conyngham, G., McNamara, G., Altrichter, H. & O'Hara, J. (2015). From Inspection to Quality: Ways in Which School Inspection Influences Change in Schools. *Studies in Educational Evaluation* 47, 47-57.

[doi:10.1016/j.stueduc.2015.07.002](https://doi.org/10.1016/j.stueduc.2015.07.002)

PFADMODELL BASIEREND AUF DEM WIRKUNGSMODELL VON SCHULINSPEKTIONEN



(Chi Quadrat = 778,04; df = 583; CFI = 0,986; TLI = 0,979; RMSEA = 0,012; 90 % CI = 0,010 - 0,014)

Wirkt Inspektion in unterschiedlichen nationalen/Governance-Kontexten unterschiedlich?

Altrichter, H. & Kemethofer, D. (2015). Does Accountability Pressure through School Inspections Promote School Improvement? *School Effectiveness and School Improvement* 26(1), 32-56.

Ehren, M.C.M., Gustafsson, J.-E., Altrichter, H., Skedsmo, G., Kemethofer, D., & Huber, S.G. (2015). Comparing effects and side effects of different school inspection systems across Europe. *Comparative Education* 51(3), 375-400.

TAB. 9: “VERMITTELNDE MECHANISMEN” NACH RECHENSCHAFTSSYSTEMEN

<i>Intermediate mechanisms</i>	<i>Accountability system</i>	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>p (≤.05)</i>
Setting expectations	Austria	310	3.34	.563	ENG > SE > AT
	Sweden	357	3.55	.689	
	England	237	3.96	.634	
Stakeholders sensitive to reports	Austria	288	3.32	.614	ENG > SE > AT
	Sweden	351	3.56	.626	
	England	238	4.03	.626	
Accepting feedback	Austria	305	3.88	.661	AT ≈ SE > ENG
	Sweden	359	3.86	.606	
	England	238	3.71	.747	

Note: The last column presents the results of the comparison of group means (Duncan test) and is to be read in the following way: “>” ... the group mean is significantly bigger ($p \leq .05$) than ...; “≈” ... the group mean does not significantly differ ($p \leq .05$) from ...

ERGEBNISSE DER FORSCHUNG ZUR REZEPTION VON FEEDBACK AUS INSPEKTIONEN UND LERNSTANDSERHEBUNGEN

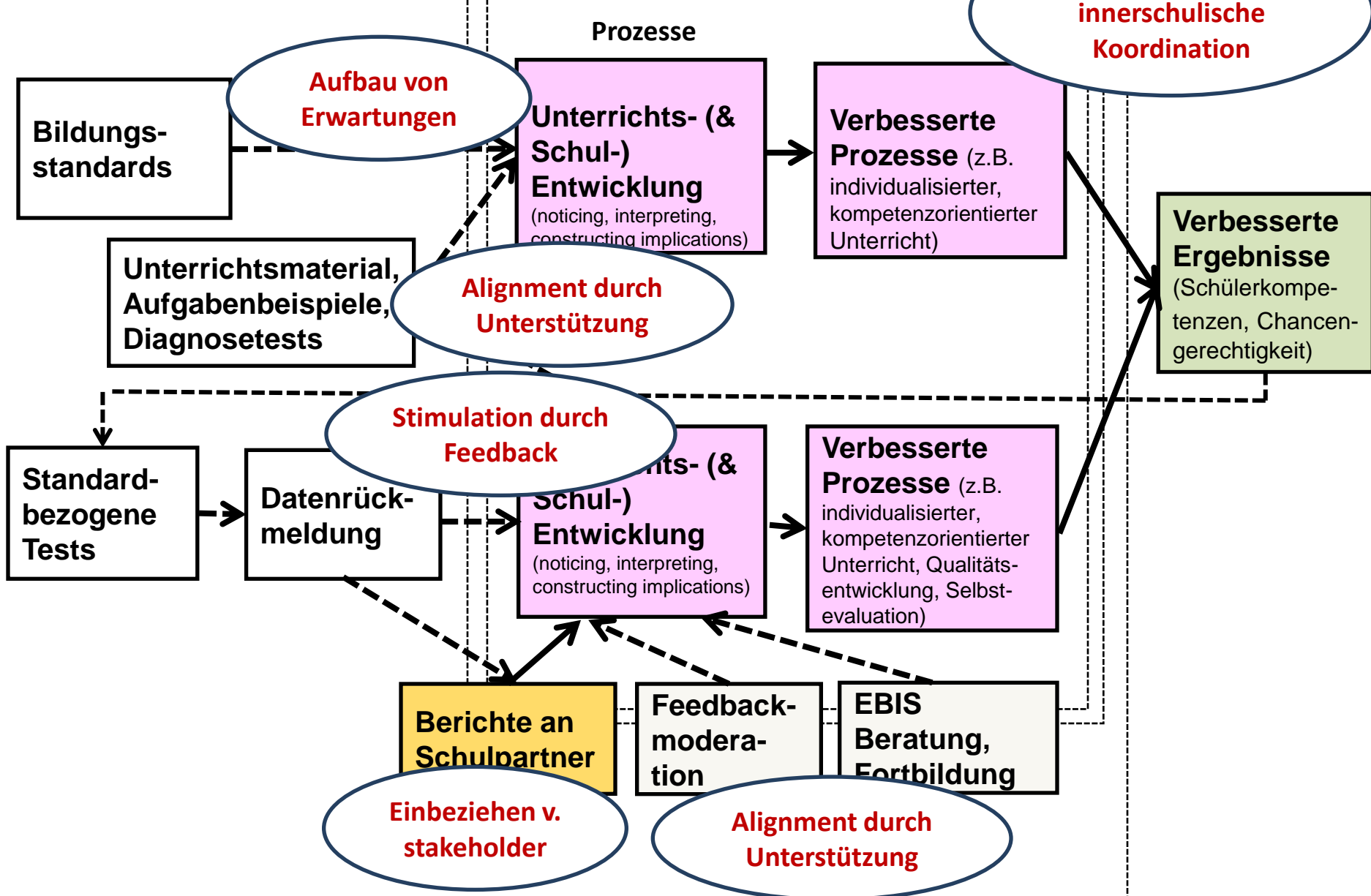
- Auch jenen Schulen und Lehrpersonen, die den neuen Steuerungsinstrumenten positiv gegenüberstehen, fällt es oft nicht leicht, aus Evaluationsrückmeldungen produktive Konsequenzen für die Weiterentwicklung von Unterricht zu ziehen
- Jene, die Konsequenzen ziehen, berichten eher „kleinere“ Veränderungen (Stoffe wiederholen, mehr Zeit für Stoffe, Testformate übernehmen)

- Mehr Rechenschaftsdruck (high stakes-Systeme)
 - vergrößert das Ausmaß an Entwicklungsaktivitäten,
 - vermindert aber die Auseinandersetzung mit Evaluationsfeedback
 - erhöht das Ausmaß von nicht-intendierten Effekten (teaching to the test, Einschränkung des Curriculums auf geprüfte Inhalte).

SCHULENTWICKLUNG DURCH BILDUNGSSTANDARDS? (STEINER ET AL. 2016)

- Frage: Rezeption von Bildungsstandards und Datenfeedback
- Fallstudien in 6 Volksschulen, Haupt-/NMS und AHS, cross case analysis
- 3 Datenerhebungen in drei Jahren:
 - vor M-Bildungsstandard-Testung, nach M-Feedback/vor E-Testung, nach E-Feedback
- Interviews mit SL, LL, Elternvertreter/in

WIRKUNGSMODELL DER 'STANDARDS-POLITIK'

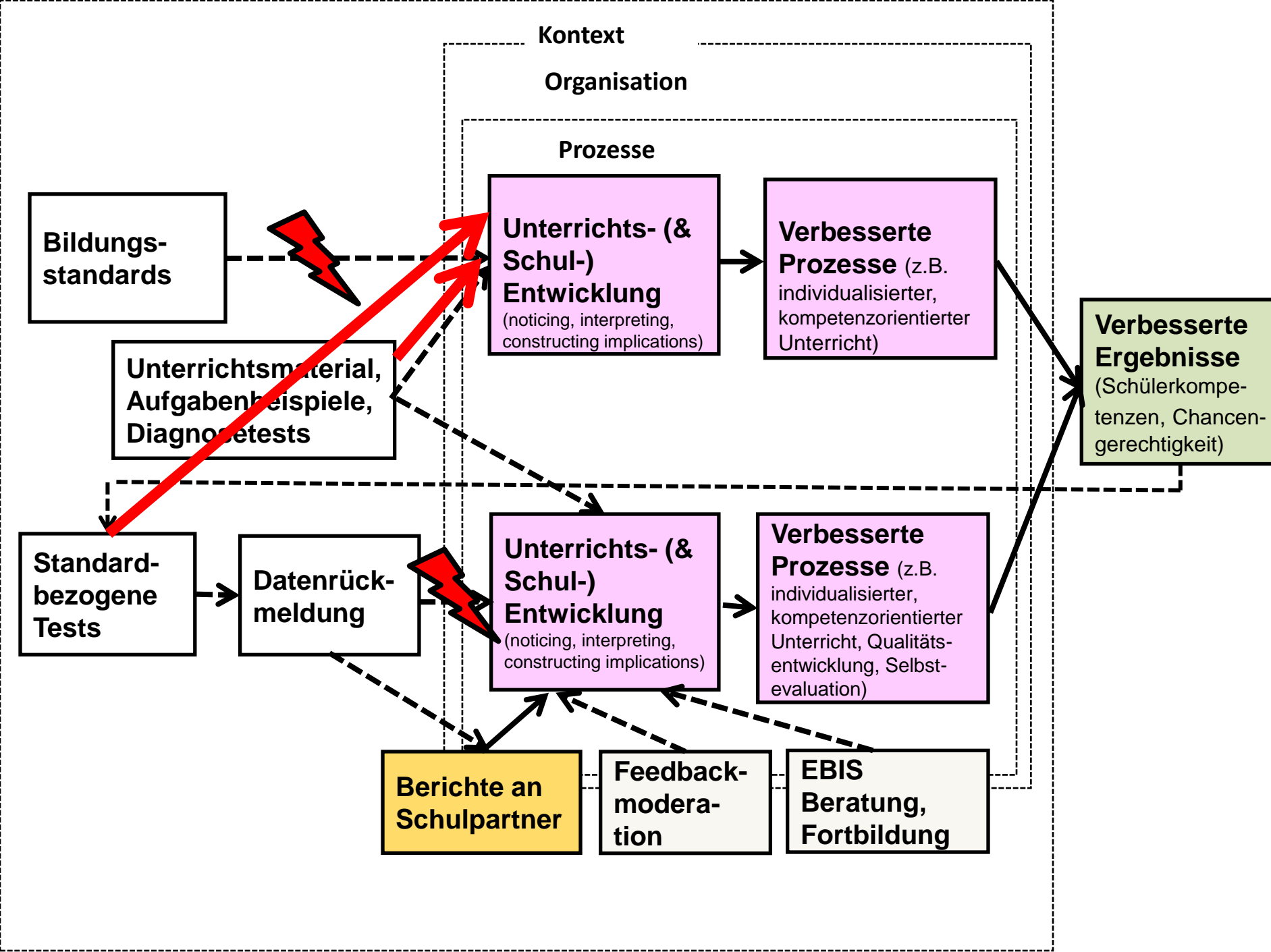


SCHULENTWICKLUNG DURCH BILDUNGSSTANDARDS? ERGEBNISSE

- Wenig Auswirkungen der Rückmeldung des Standard-Ergebnisse:
 - wenn im Erwartungswert
 - Aufgabe von Einzelpersonen, keine schulbezogene Verarbeitung
 - Keine schülerbezogenen Informationen
 - Wann komme ich wieder dran?

- Als Vorbereitung auf die Testung werden
 - in den 4. Klassen häufiger die BIFIE-Unterstützungsmaterialien studiert
 - wo notwendig, inhaltliche Umgewichtungen vorgenommen und mit neuen Aufgabentypen gearbeitet
 - Teaching to the test oder Anreicherung des Unterrichts mit neuen Lernmöglichkeiten?

- Berichte an stake holders werden nicht intensiv genutzt



ÜBERSICHT

1. 'Modernisierung' des Schulwesens und die Logik der 'evidenzbasierten Steuerung'
2. Forschung zur 'evidenzbasierten Steuerung'
 1. Schulinspektion
 2. Bildungsstandards und Datenfeedback
3. Zusammenfassung und Ausblick: Gehen QIBB und SQA darüber hinaus?

WIRKUNGSHYPOTHESEN ZUR EVIDENZBASIIERTEN STEUERUNG

- Arrangements evidenzbasierter Steuerung fördern Schulentwicklung durch den *‘Aufbau von Erwartungen’*
 - *Neo-institutionalistische Theorien: normativer Druck, Suche nach Legitimität*
- Die *‘Einbeziehung von stakeholders’* hat in unseren Schulsystemen weniger Tradition und entfaltet daher seltener stimulierende Wirkung für Unterrichts- und Schulentwicklung.
- *Datenfeedback* scheint derzeit weniger Einfluss auf Schul- und Unterrichtsentwicklung zu haben, als ursprünglich erwartet.
 - Schulleitungen nutzen Daten eher als Lehrpersonen (Young 2016)
- In der stärkeren *Abstimmung zwischen Schulleitung und Lehrpersonen* liegt möglicherweise Potential
- Unterrichtsnahe, handlungsorientierte *Unterstützungsmaterialien* fördern Veränderung

WIRKUNGSHYPOTHESEN ZUR EVIDENZBASIERTEN STEUERUNG

- Mehr Rechenschaftsdruck (high stakes-Systeme)
 - erhöht die Anpassung an vorgegebene Kriterien,
 - vergrößert das Ausmaß an Entwicklungsaktivitäten,
 - vermindert aber die Auseinandersetzung mit Evaluationsfeedback
 - erhöht das Ausmaß von nicht-intendierten Effekten (teaching to the test, Einschränkung des Curriculums auf geprüfte Inhalte).

WIRKUNGSHYPOTHESEN ZUR EVIDENZBASIIERTEN STEUERUNG

- *Neo-institutionelle Theorien* und *Theorien der sozialen Koordination* haben mehr Erklärungskraft für das Funktionieren neuer Steuerung als Feedbacktheorien
- *Feedback über eigene Leistungen* ist weniger effektiv unter high-stakes Bedingungen, weil
 - ‘gute Schulen’ sich *vor* der Inspektion/Prüfung vorbereiten ← Unter normativen Druck suchen Schulen *Legitimität* durch Konformität und Vorbereitung
 - Wenn also der Mechanismus „Aufbau von Erwartungen‘ effektiv ist, dann reduziert sich die Wahrscheinlichkeit, dass Schulleitungen “Aufmerksamkeit auf das Feedback” legen und Handlungskonsequenzen daraus ableiten
 - ‘Schulen mit Problemen’ finden es schwer, Feedback zu verarbeiten und Handlungskonsequenzen zu ziehen unter dem Handlungsdruck, der durch ungünstige Ergebnisse entstanden ist

QIBB / SQA

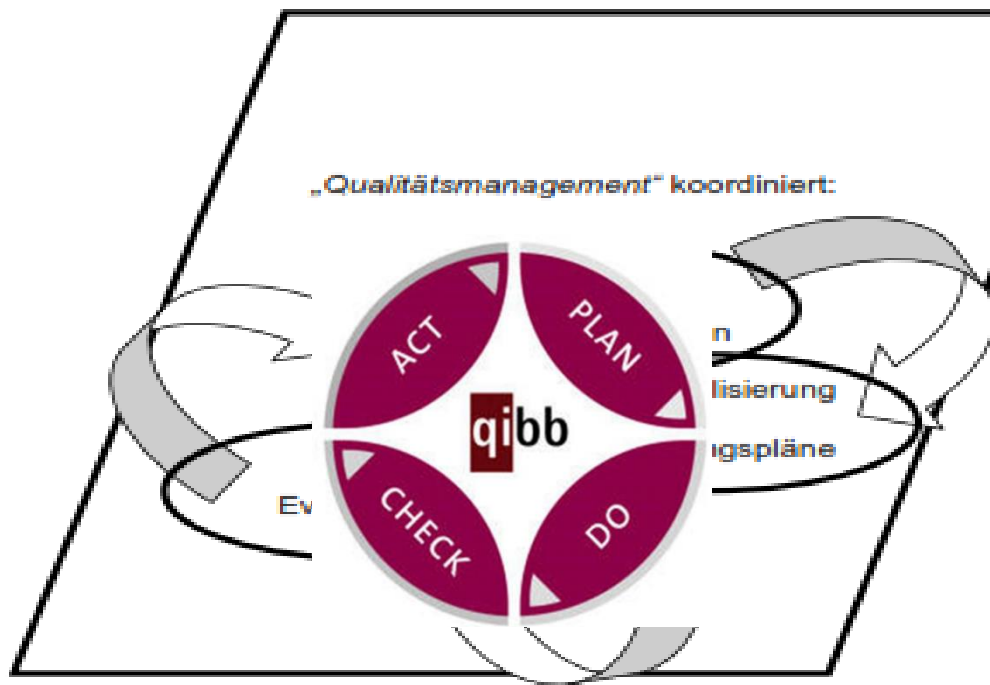
QIBB (2016); RADNITZKY (2015); MARITZEN (2015)

- Ähnlichkeiten mit herkömmlichen Instrumenten ‚evidenzbasierter Steuerung‘??
 - Wollen Schulentwicklung durch einen datenbasierten Qualitätsmanagementkreislauf ‚rationalisieren‘

Definition und Beschreibung von Schulqualität

Unterstützungsangebote, Bereitstellung von Steuerungs- und (Selbst-)Evaluierungsinstrumenten

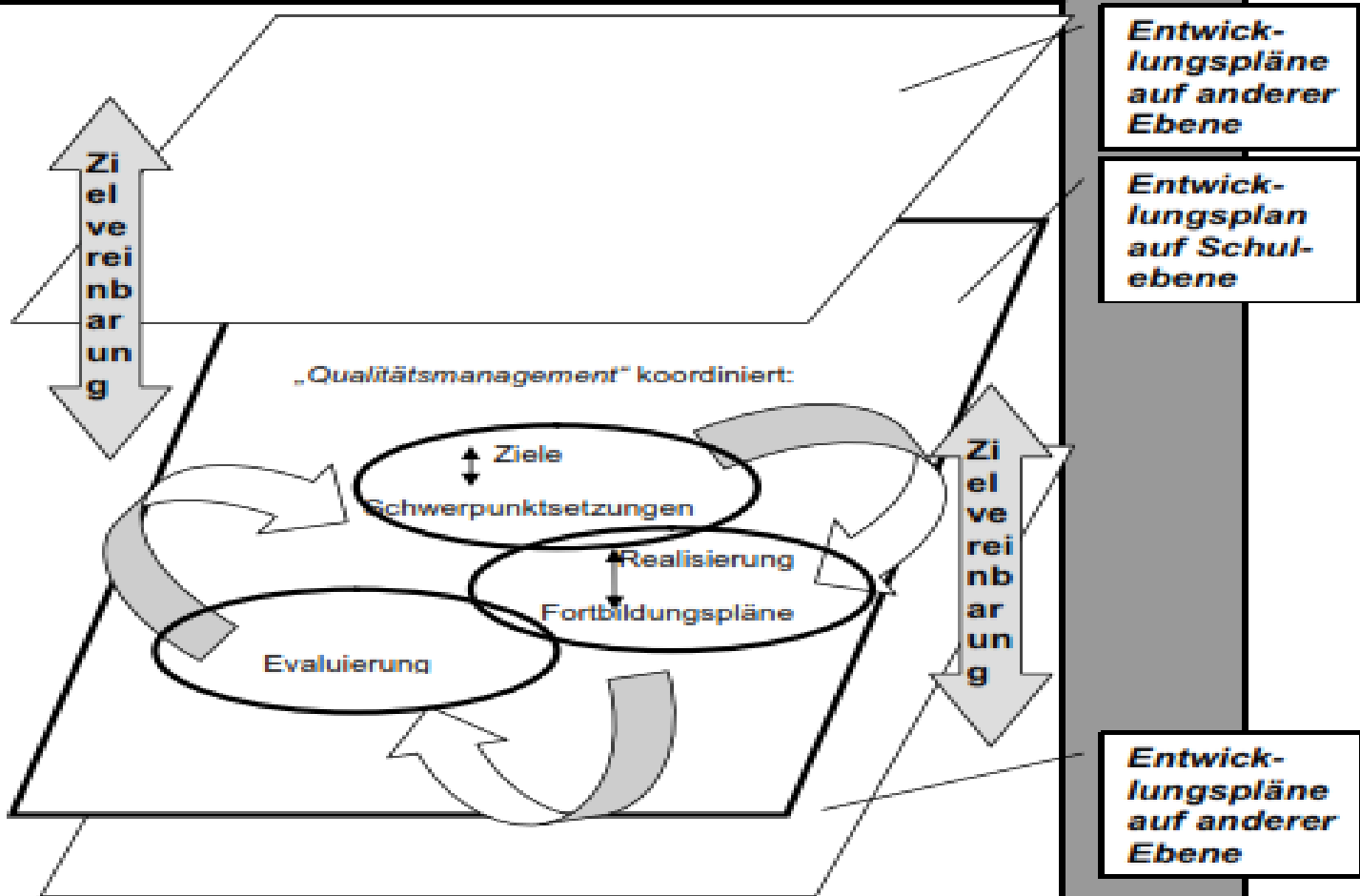
Definition und Beschreibung von Schulqualität



Entwicklungsplan auf Schulebene

Unterstützungsangebote, Bereitstellung von Steuerungs- und (Selbst-)Evaluierungsinstrumenten

Definition und Beschreibung von Schulqualität



Unterstützungsangebote, Bereitstellung von Steuerungs- und (Selbst-)Evaluierungsinstrumenten

QIBB / SQA

QIBB (2016); RADNITZKY (2015); MARITZEN (2015)

■ Ähnlichkeiten?

- Wollen Schulentwicklung durch einen datenbasierten Qualitätsmanagementkreislauf ‚rationalisieren‘

■ Unterschiede?

- Schulen nicht reaktiv auf externe Evaluation, sondern proaktiv planend und entwickelnd
- Mehrere (extern und intern produzierte) Datenquellen nutzen
- Engere, entwicklungsbezogene Bindung durch Bilanz- und Zielvereinbarungsgespräche (BZG)
- Breite Unterstützungsumgebung
- Konsequenterer Umsetzung vergleichbarer Prozesse auf allen Ebenen

- „... neue Generation von Qualitätspolitikern, die ältere Konzepte der Einzelschulentwicklung (wie sie sich beispielsweise im Schulprogramm und in der einzelschulischen Festlegung von Qualitätszielen ausdrücken) mit der Nutzung der neuen Instrumente evidenzbasierter Steuerung verbinden wollen.“
- „Anerkennung und Wertschätzung von Lehrerprofessionalität“

WAS TUN, UM DATENNUTZUNG ZUR SCHUL- UND UNTERRICHTS-ENTWICKLUNG ZU FÖRDERN?

- Bessere Rückmeldungen?
- Qualifikation von Schulleitungen und Lehrpersonen?
 - Hängt mangelnde Nutzung von Datenfeedback mit mangelndem Wissen über Evaluationsmethoden zusammen?
 - weniger Technik der Interpretation, als Schritte von Interpretation zu Unterrichtsplanung und -durchführung
- Druck erhöhen?
- Bindung zwischen Schulleitungen und Lehrpersonen?
 - Unterschiedliche Nutzung von Informationen auf verschiedenen Ebenen (Young 2016)
 - Fachgruppen als vermittelndes Element
 - SL: Verantwortung und Beispiele für sinnvolle Nutzung von Daten sichtbar machen
 - ‚Feedbackkultur‘? – längerfristige Signale aus der Umgebung
- Unterstützung durch Materialien und handlungsnahe Unterstützungssysteme:

Herzlichen Dank



„UMSTEUERUNG“ – WOVON UND WIE?

- Seit dem PISA-Schock ist Qualitätsevaluation nicht nur ein Hilfsmittel zur ‚rationalen‘ Einzelschulentwicklung, sondern ein wesentliches Element einer „Umsteuerung“ des Bildungswesens.
- Eine Grundidee dieser „Umsteuerung“ („Evidenzbasierte Steuerung“) besteht darin, dass Qualitätsentwicklung zielgerichteter, auf einer besseren Informationsbasis und schneller als bis dahin üblich betrieben werden müsse und dass dafür nicht nur einzelschulische, sondern auch gesamt-systemische Massnahmen notwendig wären.
- Besonders wichtig sind «Output-Informationen» über Schülerleistungen

UNTER WELCHEN BEDINGUNGEN KANN FEEDBACK VERARBEITET WERDEN? (KLUGER & DENISI 1996; HATTIE UND TIMPERLEY 2007)

Wirksamkeit von Feedback hängt ab von:

- *Cues im Feedback*: Aufmerksamkeit auf *Aufgabenbearbeitung* und *Selbstregulation* lenken
- *Charakteristika der Aufgabe*: eher bei einfachen Aufgaben
- *Situations- und persönlichkeitsbezogene Variablen*: Klare, spezifische und herausfordernde Ziele; geringe Bedrohung des Selbstwertes

UNTER WELCHEN BEDINGUNGEN KANN FEEDBACK VERARBEITET WERDEN? (KLUGER & DENISI 1996; HATTIE UND TIMPERLEY 2007

- Feedbackmechanismus wirkt auch im interpersonellen Bereich nur dort, wo das Repertoire vorhanden und leichte Verhaltensanpassung notwendig ist
- wo grundlegende Umsteuerung notwendig ist, braucht es andere Maßnahmen
- ohne Bedrohung des Selbstwerts
- klare und bewältigbare „Lernziele“ der Lehrpersonen
- bei der Formulierung der Feedbackfragen/dimensionen zu beteiligen
- Information über Aspekte der Aufgabenbewältigung (= Unterrichtsmerkmale)

SCHÜLERLEISTUNG UND UNTERRICHTSENTWICKLUNG

- Annahme von Lernstandserhebungen: schülerbezogene Leistungsinformationen lassen sich – mit Hilfe von Kompetenzmodellen - in didaktische Erneuerung übersetzen

→ Leistungsinformation = «Indikator»

→ Für Unterrichtsentwicklung braucht es mehr prozessbezogene Information, mehr Information über Faktoren der Unterrichtsqualität und didaktischen Gestaltung