



UNIVERSITÄT  
DUISBURG  
ESSEN

*Offen im Denken*

## *Zwischen den Fronten?!*

*Selektion, Rechenschaftslegung und datengestützte Schulentwicklung im Kontext zentraler Abschlussprüfungen – ein internationaler Vergleich*

Esther Dominique Klein

Isabell van Ackeren

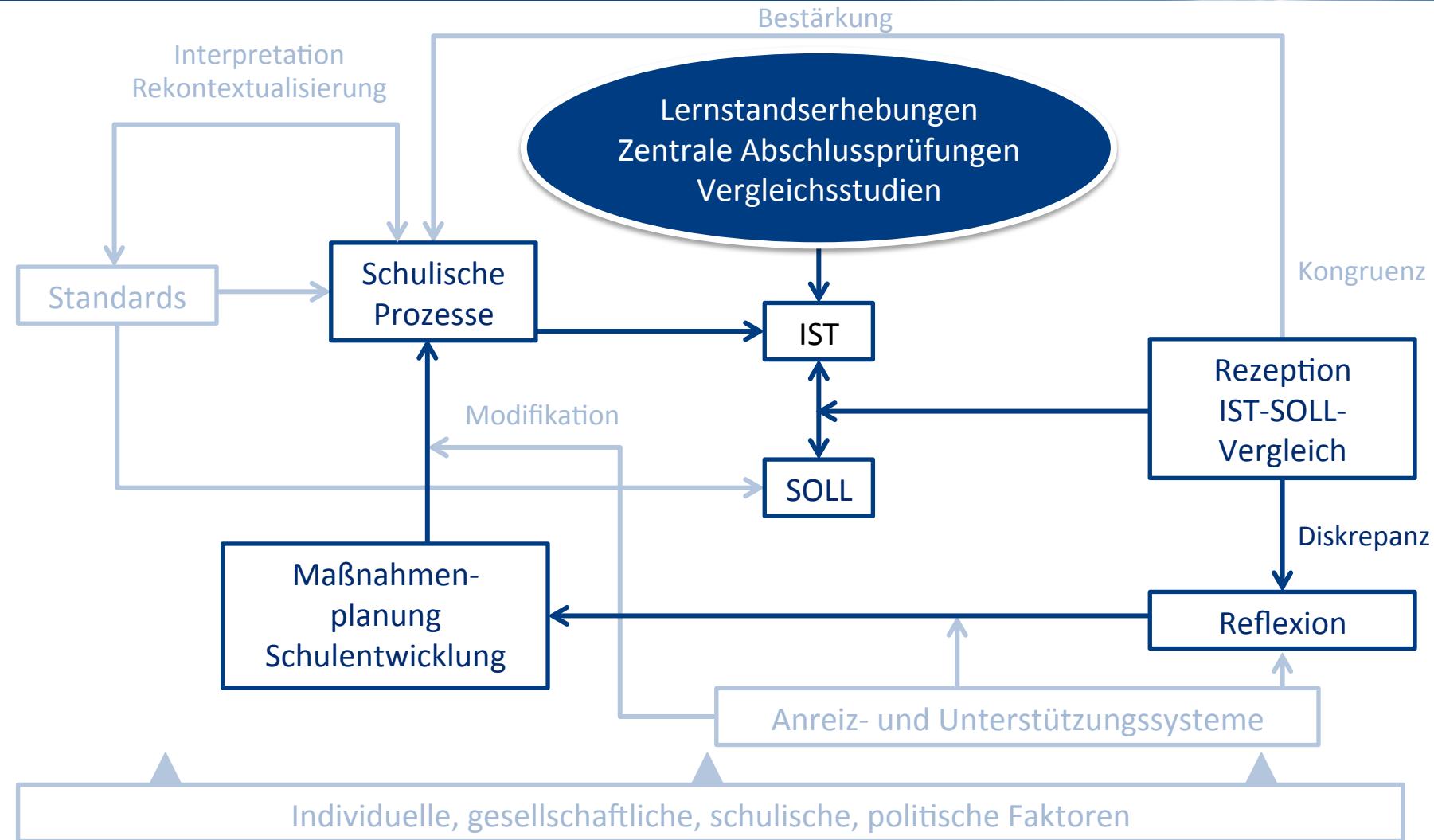
24. Kongress der DGfE – 12. März 2014

- Zentrale Abschlussprüfungen werden als einheitliches Konstrukt wahrgenommen
  - Heterogene Prüfungssysteme (*vgl. Klein & van Ackeren, 2011; Klein et al., 2009*) werden kaum thematisiert
  - Befundlage zu Leistungseffekten ist inkonsistent (*Baumert & Watermann, 2000; Bishop, 1995; Block et al., 2011; Cosentino de Cohen, 2010; Holme et al., 2010; Maag Merki, 2012; Shuster, 2012*)
- Forschung fokussiert auf Schülerleistungen und Unterrichtshandeln
  - Prozesse und Strategien im Umgang mit den Prüfungen werden kaum beleuchtet (*vgl. Klein, 2013; Klein et al., im Erscheinen*)

# Zentrale Abschlussprüfungen als Steuerungs- und Schulentwicklungsinstrument

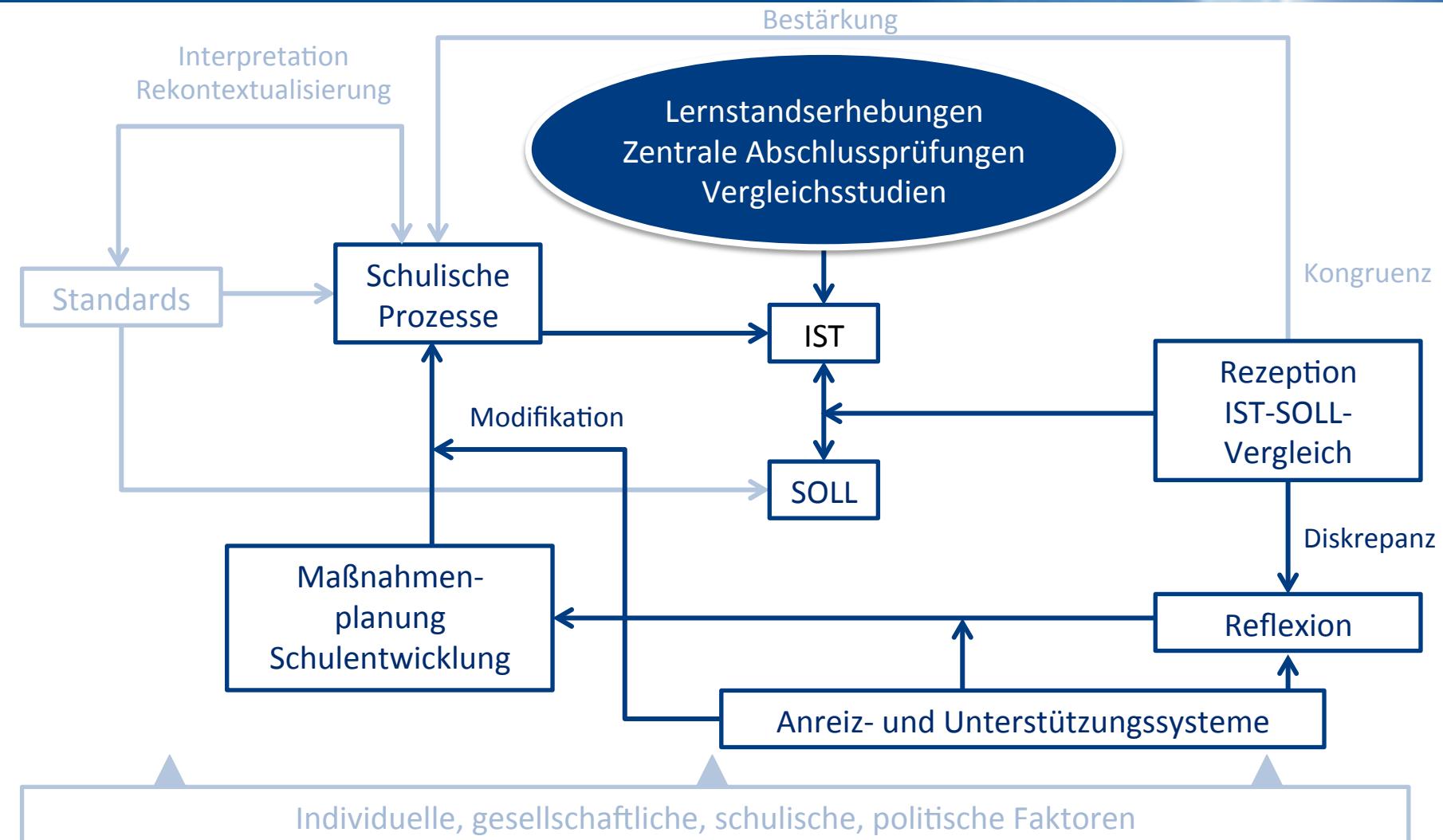
# Vermutete Wirkung von ZAP und Co.

(Maag Merki, 2010; gekürzte Darstellung)



# Vermutete Wirkung von ZAP und Co.

(Maag Merki, 2010; gekürzte Darstellung)



# Forschungsstand: Wirkungen und Nutzung von ZAP

## Prüfungsverfahren und Feedback

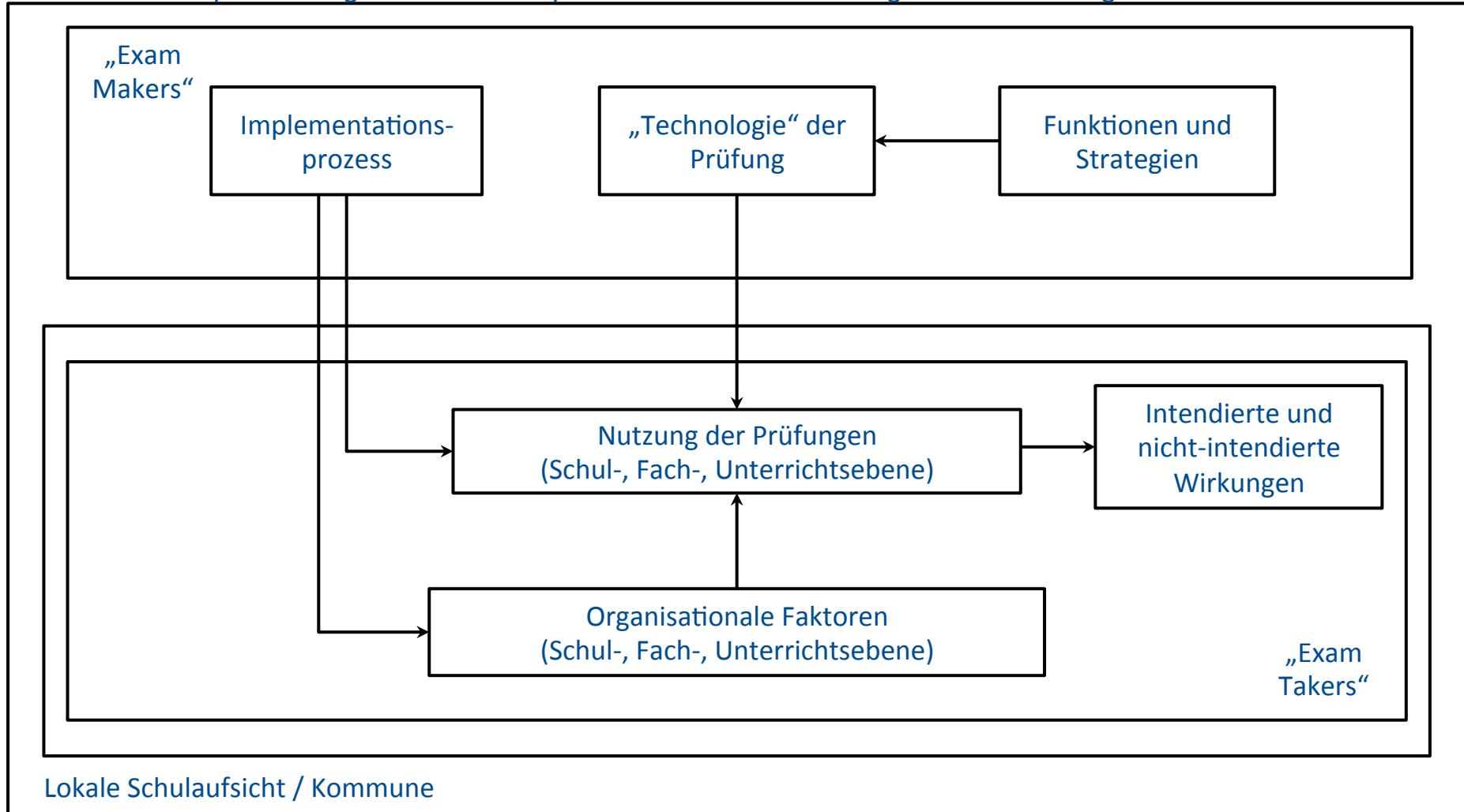
- Konsequenzen für Schüler/innen haben nur teilweise höhere Relevanz für Lehrkräfte als School Accountability; insgesamt inkonsistent (*Goertz & Massell, 2005; Massell et al., 2005; Sipple et al., 2004; Vogler, 2006, 2008; Vogler & Carnes, 2009; Zhang, 2009*)
- Wahrgenommene Nützlichkeit des retroaktiven Feedbacks eher gering (*Saunders & Rudd, 1999*)

## Schulischer Umsetzungsprozess

- Schulorganisationale Faktoren führen zu differenziellem Umgang mit Abschlussprüfungen (*DeBray, 2005; DeBray et al., 2000; Saunders, 2000; Saunders & Rudd, 1999*)
- Datennutzung und Umsetzung der Prüfungen im Unterricht fachabhängig; tendenziell Hinweise auf geringere Flexibilität (*van Ackeren et al., 2012; Eickelmann et al., 2011; Kühn & Racherbäumer 2013; Maag Merki, 2012; Tymms, 1995; Vogler, 2006, 2008; Vogler & Carnes, 2009*)
- Notwendigkeit der Unterstützung von Schulentwicklungsprozessen (*DeBray et al., 2000*)
- Geringe Nutzung von Daten trotz Training (*Tymms, 1995*)

# Ein Modell zur Analyse der Nutzung zentraler Abschlussprüfungen

Governance-System und grundsätzliches System der Qualitätssicherung und -entwicklung



# Anlage der internationalen Vergleichsstudie

## Teilstudie 1:

### **Zentrale Abschlussprüfungen als Steuerungsinstrument**

- Welche Funktionen sollen zentrale Abschlussprüfungen in verschiedenen Kontexten erfüllen?
- Wie werden diese Funktionen von schulischen Akteuren wahrgenommen?

## Teilstudie 2:

### **Zentrale Abschlussprüfungen als SchulentwicklungsInstrument**

- Auf welche Weise nutzen schulische Akteure die Prüfungen für ihre Arbeit?
- Durch welche Faktoren wird die Art und Weise der Nutzung beeinflusst?
- Inwiefern können zentrale Abschlussprüfungen die Arbeit in Schulen lenken?

- International vergleichende Fallstudie zentraler Abschlussprüfungen am Ende der Sekundarstufe II
- Multimethodaler, mehrperspektivischer Ansatz (*Yin, 2009*)

## **Teilstudie 1: Funktionen der Prüfungen**

- Experteninterviews mit Vertreter/innen der Prüfungsbehörden ( $N = 9$ )
- Leitfadengestützt
- Inhaltsanalytische Auswertung
- Spiegelung mit Wahrnehmung der schulischen Akteure (siehe Teilstudie 2)



## **Teilstudie 2: Schulische Nutzung der Prüfungen**

- Standardisierte Befragung mittels Fragebogen
- $N = 35$  Schulleitungen
- $N = 385$  Lehrkräfte

- Exploratives Vorgehen

Kontextfreie Forschungsfrage → Deskription → Juxtaposition und Vergleich → Generierung von Hypothesen (*Philips & Schweisfurth, 2007*)

# Charakteristika der ausgewählten Fälle

	Finnland	Irland	Niederlande
Prüfung	<i>Ylioppilastutkinto</i>	<i>Leaving Certificate Examinations</i>	<i>Eindexamen vwo</i>
Prüfungsart	Zentrale schriftliche Prüfungen	Zentrale schriftliche Prüfungen	Zentrale schriftliche (50%) und dezentrale Prüfungen (50%)
Standardisierung des Prüfungsverfahrens	moderat	hoch	Zentrale Prüfungen: hoch Dezentrale Prüfungen: niedrig
Governancesystem*	Legitimitätstyp mit <i>Local Empowerment</i>	<i>School Empowerment</i> mit Legitimitäts- und Bürokratiertyp	Legitimitäts- und Effizienztyp mit <i>School Empowerment</i>

\*vgl. Schmid et al., 2007; Glatter et al., 2003

# Charakteristika der Prüfungsverfahren

## Finnland

- Formale Qualifikation der Schüler/innen
- Kein Versuch, Schulen zu beeinflussen
- Keine Inspektion
- Datenrückmeldung als Unterstützungsangebot
- Keine Entwicklungsforderung

## Irland

- Formale Qualifikation der Schüler/innen im Vordergrund
- Vollständig extern
- Keine Kontrolle von Schulen durch Prüfungen
- Niederschwelliges Inspektionsverfahren
- Keine Entwicklungs-erwartungen

## Niederlande

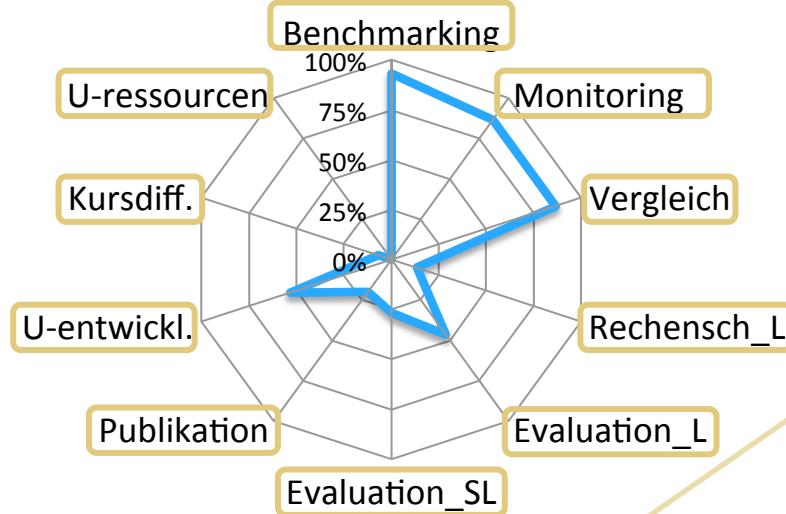
- 50/50-System
- Prüfungsergebnisse als Benchmarking von Schüler- und Schulleistungen
- Mehrdimensionales internes und externes Evaluationssystem
- Klare Entwicklungs-forderungen an Schulen

# Nutzung der Prüfungen auf Schul-, Fach- und Individualebene

# Umfang der Nutzung auf Schulebene

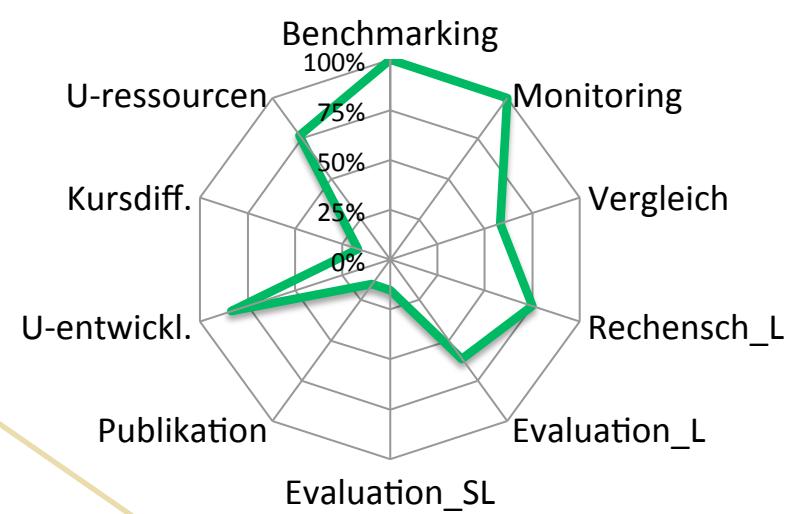
Finnland

N = 15

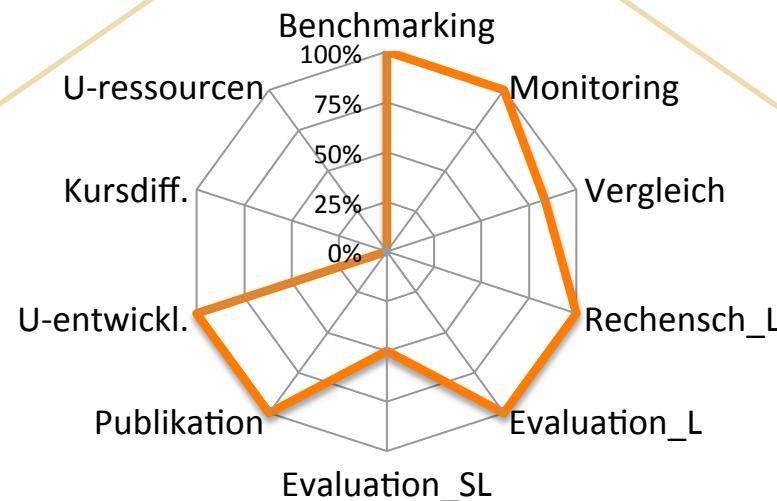


Irland

N = 12



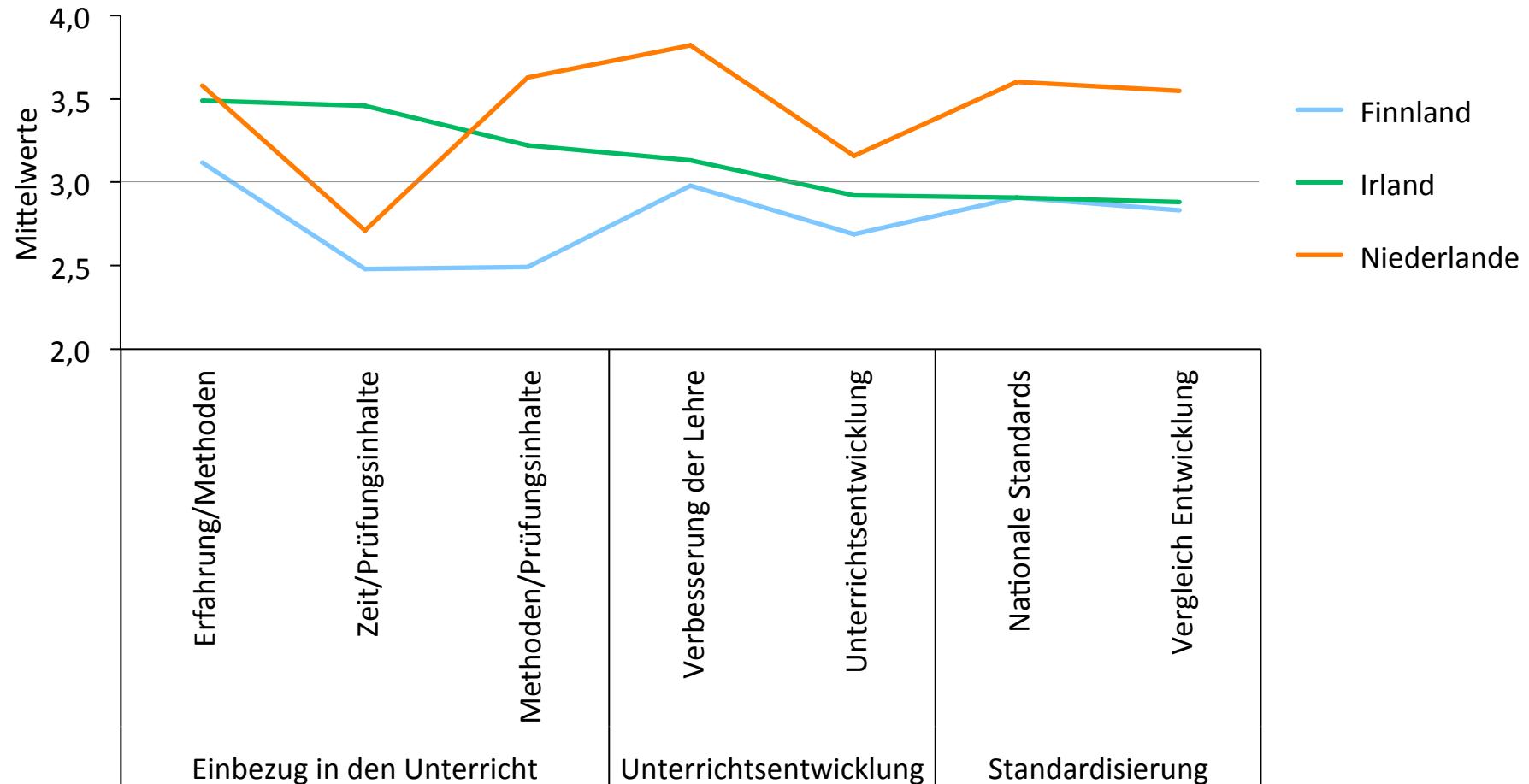
Niederlande



N = 6

E. Dominique Klein  
Isabell van Ackeren

# Fachebene: Kooperation im Rahmen der Prüfungen

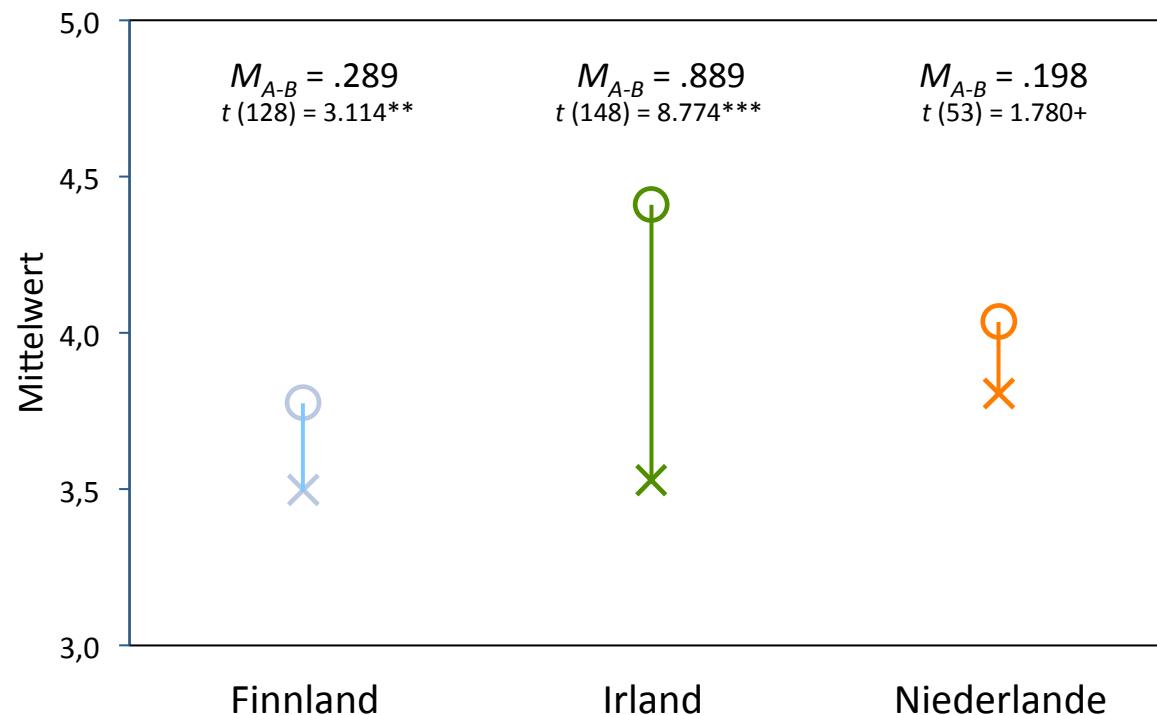


5-stufige Likertskala; 1 = trifft überhaupt nicht zu; 5 = trifft voll und ganz zu

# Individualebene: Nutzung der Prüfungsaufgaben und -ergebnisse

*Nutzung der Prüfungsaufgaben/-ergebnisse für die...*

✗ Unterrichtsentwicklung    ○ Leistungsbewertung

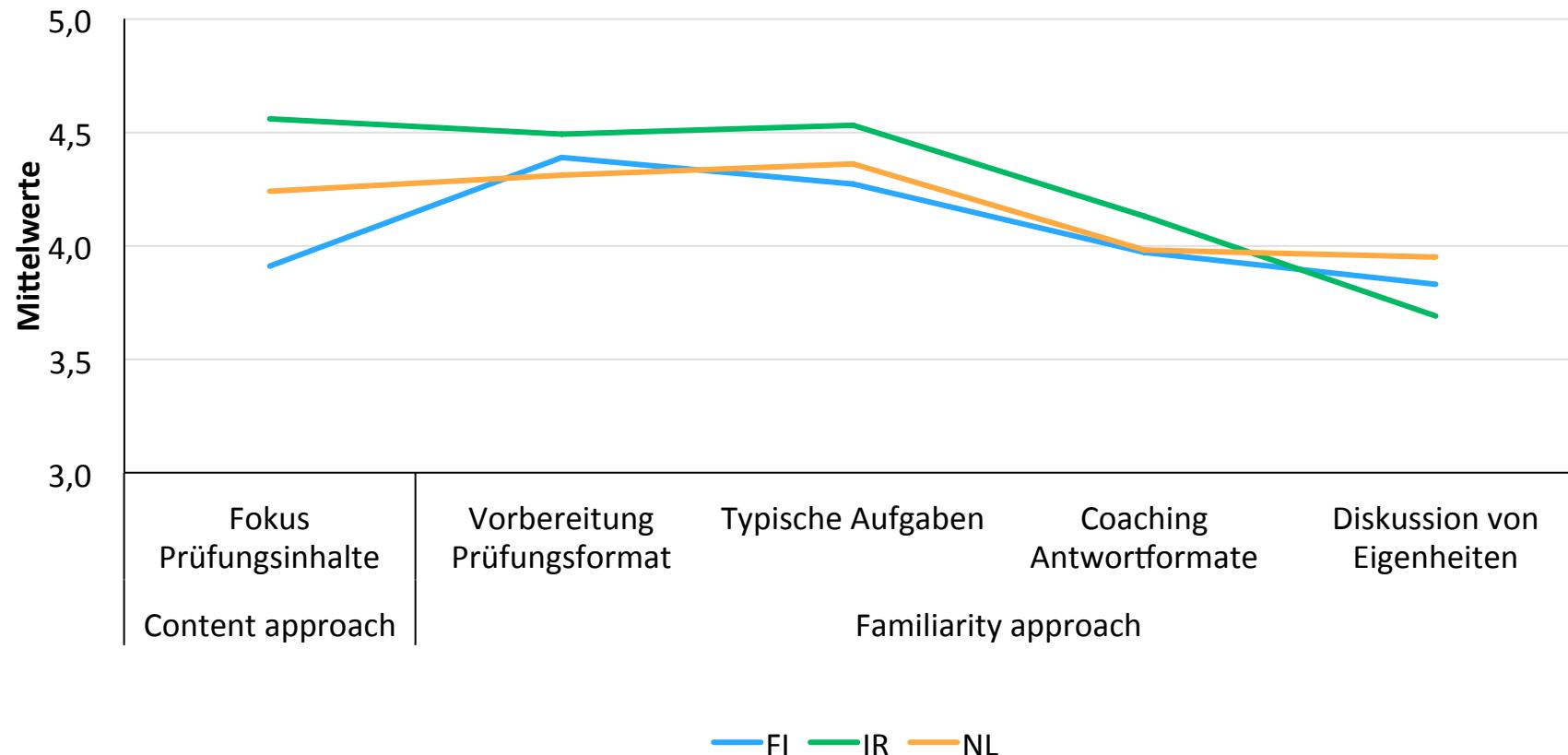


5-stufige Likertskaala; 1 = stimme überhaupt nicht zu; 5 = stimme voll und ganz zu

\*\*\*  $p < .001$ ; \*\*  $p < .01$ ; \*  $p < .05$ ; +  $p < .1$

# Vorbereitungsstrategien der Lehrkräfte

# Individualebene: Vorbereitungsstrategien

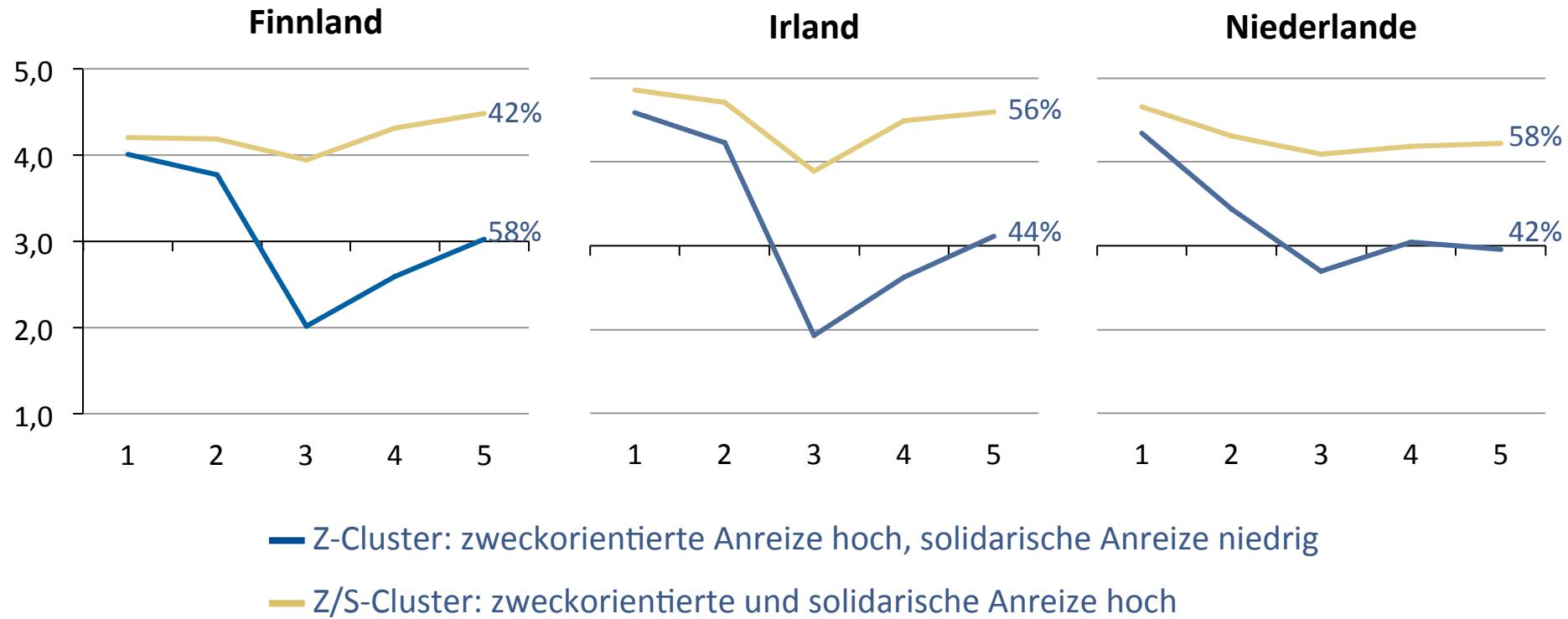


5-stufige Likert-Skala; 1 = nie – 5 = sehr oft

# Zusammenhang zwischen Prüfungsanreizen und Nutzung der Prüfungen

# Motivation durch die Prüfungen

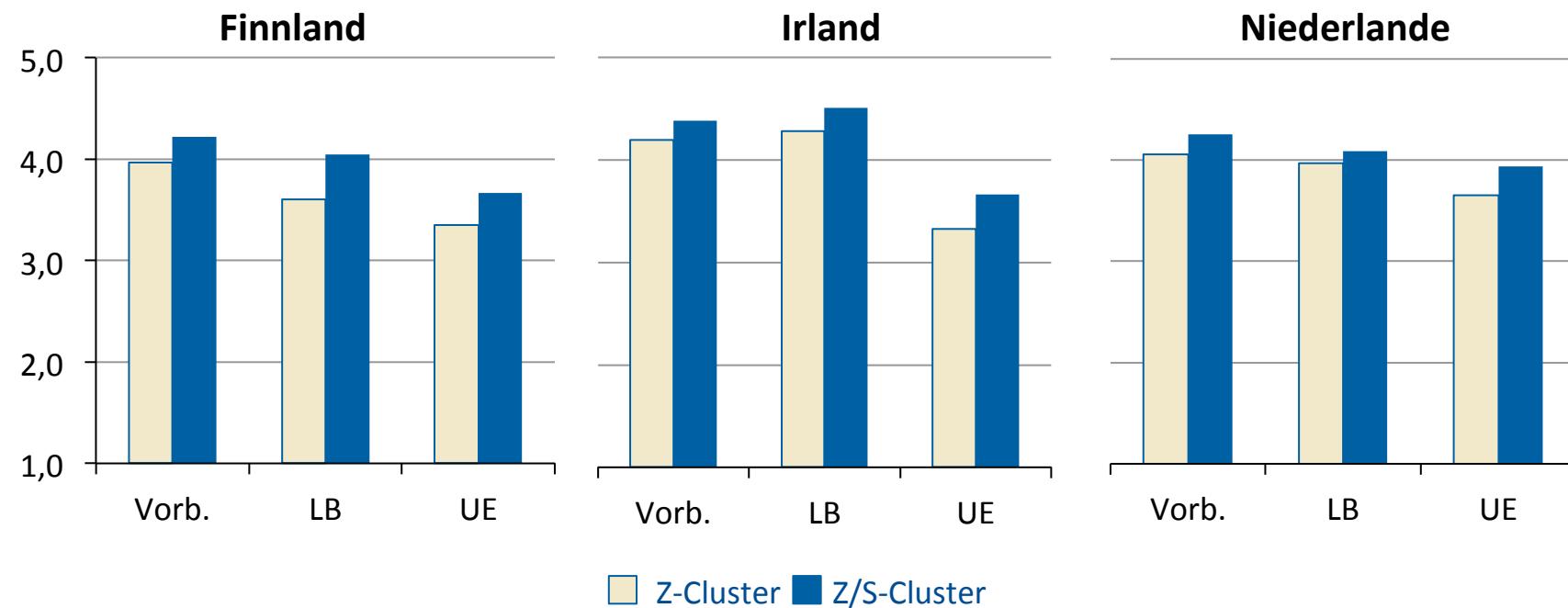
(hierarchische Clusteranalyse: Ward & quadrierte euklidische Distanz; Optimierung: k-Means)



„Ich bin motiviert, meine Schüler/innen optimal auf die Prüfungen vorzubereiten, weil...“:  
1 = Wohlergehen SuS; 2 = Perspektiven SuS; 3 = Rechenschaft gegenüber SL; 4 = Profilierung  
Schule; 5 = Attraktivität Schule

5-stufige Likertskala (1 = stimme überhaupt nicht zu – 5 = stimme voll und ganz zu)

# Motivation und Strategien im Umgang mit den ZAP



Vorb. = Vorbereitungsstrategien; LB = Leistungsbeurteilung; UE = Unterrichtsentwicklung

5-stufige Likertskaala; 1 = stimme überhaupt nicht zu; 5 = stimme voll und ganz zu

## Leistungsbewertung:

signifikanter Haupteffekt der Cluster in IR,  $F(1, 138) = 4.078^*$

signifikanter Interaktionseffekt Cluster\*Fach\*Dienstalter in FI,  $F(6, 106) = 2.205^*$

→ Keine weiteren Haupt-/Interaktionseffekte für die Clusterzugehörigkeit

	Finnland	Irland	Niederlande
Funktion der Prüfung	Formale Qualifikation & Entwicklung	Formale Qualifikation	Formale Qualifikation, Motivation & Entwicklung
Relevanz der Prüfung	Niedriger	Höher	Höher
Schulische Nutzung	Individualisierte Routine	Eindimensional zweckorientiert	Mehrdimensional systematisch

# Eine Hypothese zur Nutzung zentraler Abschlussprüfungen...

**Durch die Konsequenzen für die Schüler/innen wird die Aufmerksamkeit der Lehrkräfte auf die Leistungen der Schüler/innen, nicht aber auf regulative Strategien im Hinblick auf die Qualität des Unterrichts gelenkt.**



Zentrale Abschlussprüfungen werden zur Schul- und Unterrichtsentwicklung genutzt, wenn es gelingt, die Aufmerksamkeit auf den **Zusammenhang zwischen Prozess- und Produktqualität** zu lenken:

1. Auf Schulentwicklung gerichtete Kooperationsstrukturen
2. Gesamtsystem schulischer Qualitätsentwicklung

***Vielen Dank für Ihre  
Aufmerksamkeit!***

*Esther Dominique Klein  
Isabell van Ackeren*

*Universität Duisburg-Essen  
Fakultät für Bildungswissenschaften*

*Kontakt:*

*[dominique.klein@uni-due.de](mailto:dominique.klein@uni-due.de)*