



Wie rezipieren Lehrkräfte (quantitative) Rückmeldungen ihrer Unterrichtsqualität?

Samuel Merk - Universität Tübingen

Übersicht

- Datenbasierte Unterrichtsentwicklung (in Deutschland)
- Forschungsstränge zur datenbasierten Unterrichtsentwicklung
- Der Rezeptionsprozess im Fokus
- Die vorliegende Studie
 - Motivation
 - Design, Material und Analysen
 - Ergebnisse
- Diskussion



Datenbasierte Unterrichtsentwicklung

Datenbasierte Unterrichtsentwicklung (in Deutschland)

- Rahmenkonzeption der KMK (2006) zum Bildungsmonitoring
- Differentielle Datenquellen und -adressaten

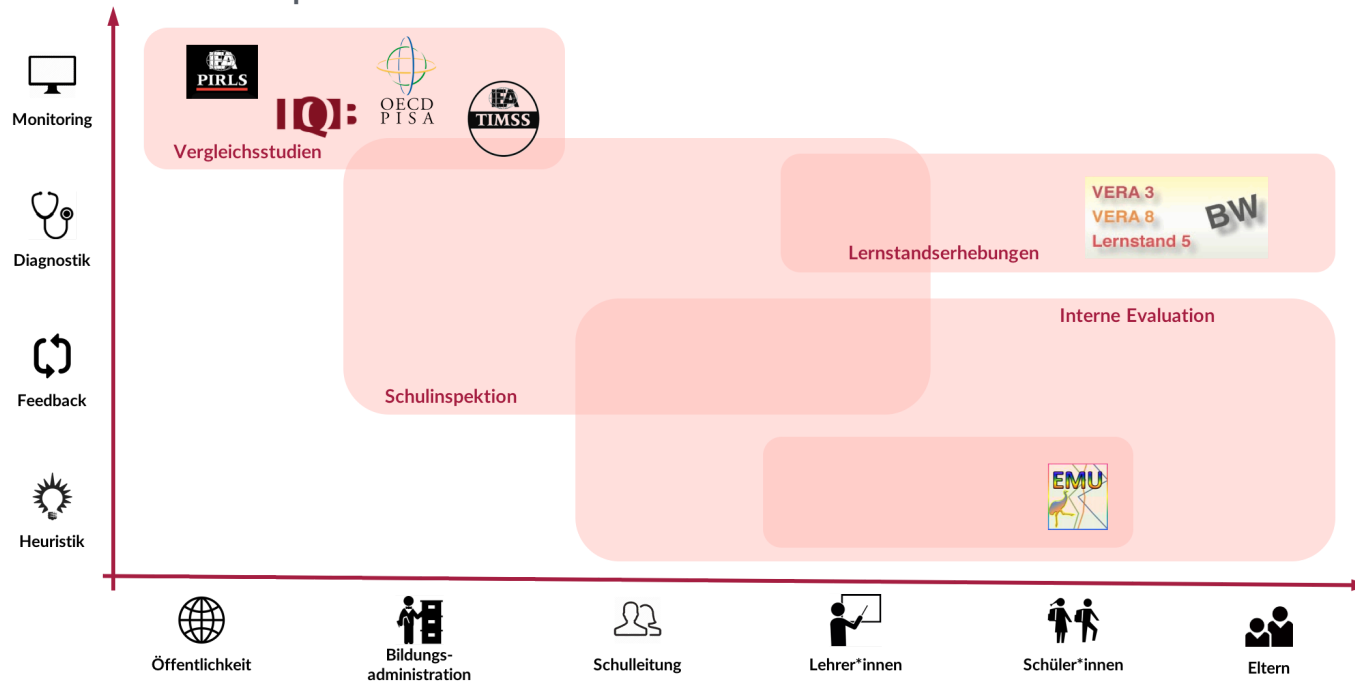


Abb. 1: Funktionen und Adressaten

Datenbasierte Unterrichtsentwicklung (in Deutschland)

- Rahmenkonzeption der KMK (2006) zum Bildungsmonitoring
- Differentielle Datenquellen und -adressaten

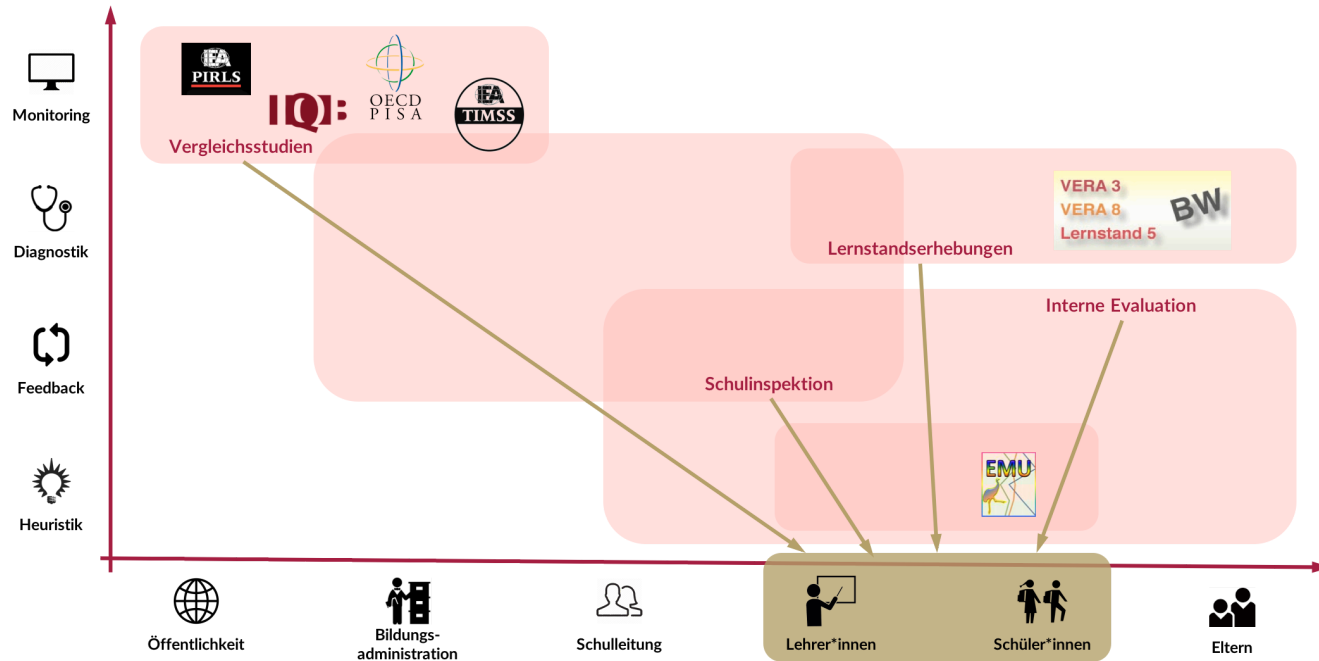


Abb. 1: Funktionen und Adressaten

Forschungsstränge zur datenbasierten Unterrichtsentwicklung

Forschungsstrang 1: Konzeptuelle Forschung

Beispiele:

- Entwicklung von prozess- (z.B. Helmke & Hosenfeld, 2005) und kontextorientierte Modellen (Verhaeghe, Vanhoof, Valcke, & Van Petegem, 2010; Visscher, 2002)

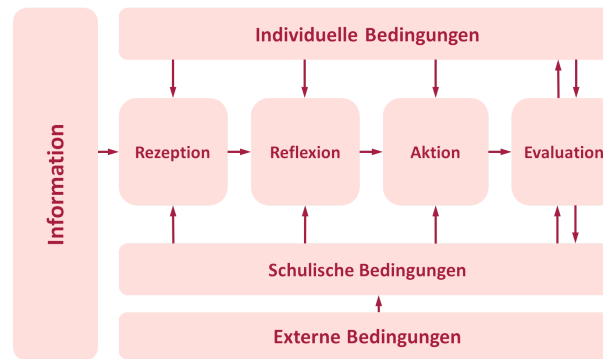


Abb. 2: Zyklusmodell "Von der Evaluation zur Innovation". Abbildung nach Helmke und Hosenfeld (2005)

- Diskussion differentieller Policies (Maier, 2010)
- Theoretische Verhältnisbestimmung diverser Verfahren (Gärtner, 2013)

Forschungsstrang 2: Deskriptive & Explanative Nutzungs- und Effectivenessforschung

Beispiele:

- Nutzung diverser Verfahren (*Maier, 2008*)
- Differentielle Wirksamkeit verschiedener Daten (*Wurster, Richter, & Lenski, 2017*)
- Kontexteffekte auf die Nutzung (*Gärtner, 2016*)



Forschungsstrang 3: Interventionsforschung

Beispiele

- Statistical Literacy Trainings (*Koch, 2011*)
- Datennutzungstrainings (*Mandinach & Gummer, 2016*)



Im Folgenden präsentierte Studien

- *Fokus* auf dem Rezeptionsprozess
 - da *notwendige Bedingung* für Reflektion und Aktion (*Helmke & Hosenfeld, 2005*)
 - da (zumindest wahrgenommenes) Ungleichgewicht zwischen Aufwand bei der Datengenerierung und Aufwand für die Implementation(-sforschung)
- *Ziel*: Entwicklung von
 - Rückmeldeformate, welche ohne statistische Vorkenntnisse (intuitiv) korrekt und motivationsförderlich rezipiert werden
 - Interventionen zur **Förderung der Datenkompetenz**



Intuitiv korrekte und motivationsförderliche Rückmeldung?

Selbstversuch

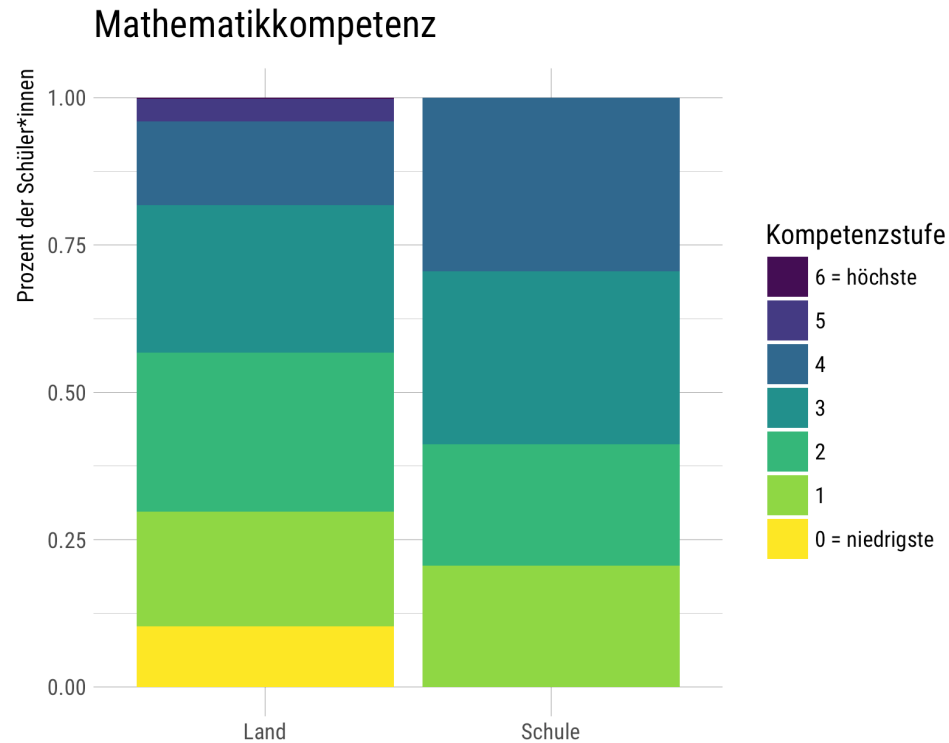


Abb. 3: Stacked Bar Graph (PISA-Daten)



Welche Graphiken sind intuitiv und motivationsförderlich?

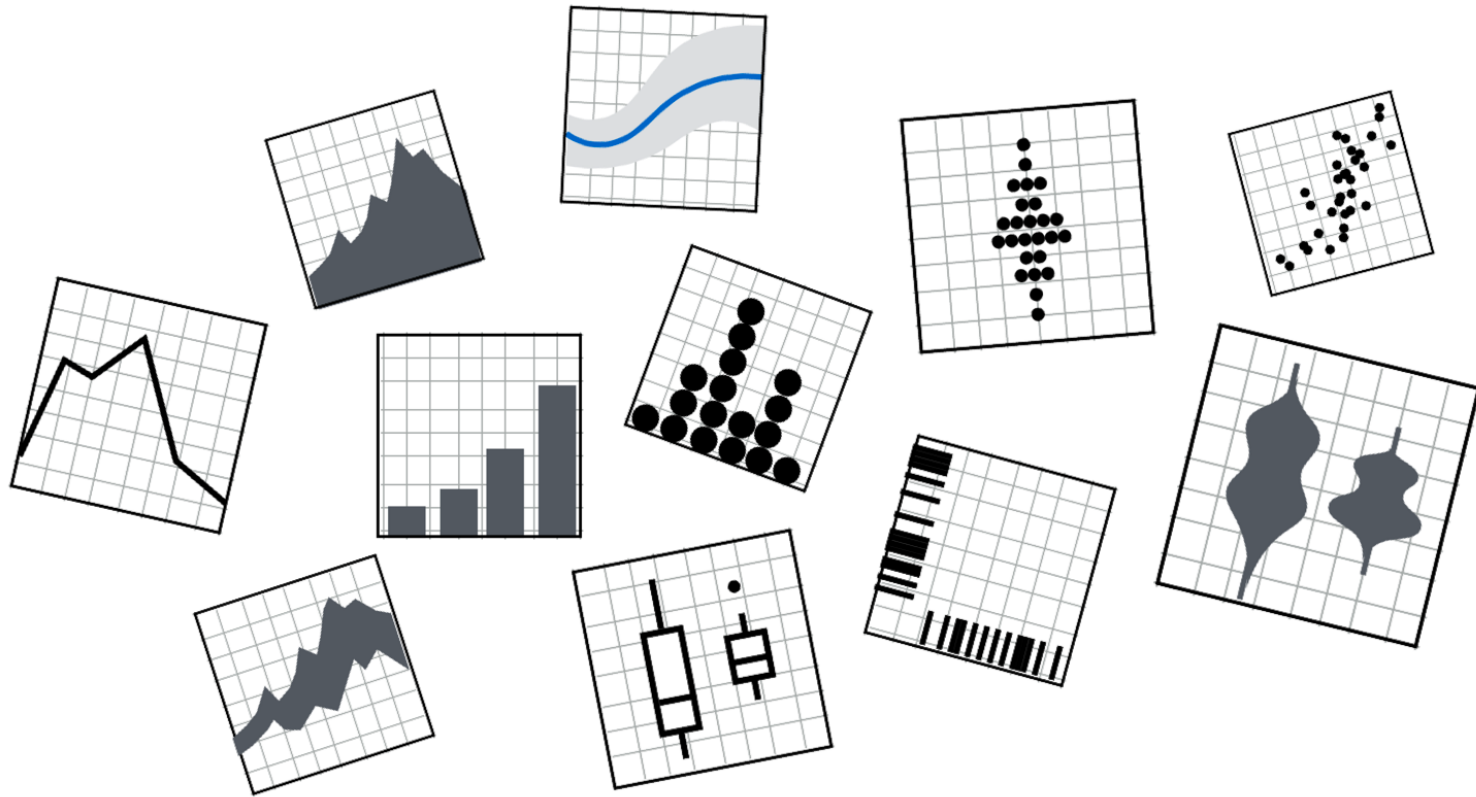


Abb. 4: Diverse Graphiken

Vorliegende Studie I

Durchführung im Rahmen eines bestehenden Qualitätssicherungs- und Rückmeldesystems einer Nachhilfeschule, welche Abiturvorbereitungskurse in Mathematik anbietet.

- Design Based Research (*The Design-Based Research Collective, 2003*) \Rightarrow Design = Kompromiss aus Forscher- und Anwenderanliegen ^{hier} = **Interaktive Onlineapplikation**



Vorliegende Studie

Forschungsfrage 1:

Welche

- **Daten** (*Single-Items, Skalenwerte, Freitextantworten*),
- welche **Datentransformationen** (*Rohwerte, z- oder ipsativ standardisierte Werte*) und
- welche **statistischen Größen**/graphischen Darstellungen (*Boxplots, Mittelwerte, CI, Streuung*)

bevorzugen Lehrkräfte bei der Rezeption von Daten aus der Unterrichtsdiagnostik?



Vorliegende Studie

Forschungsfrage 2:

Welche

- **Daten** (*Single-Items, Skalenwerte, Freitextantworten*),
- welche **Datentransformationen** (*Rohwerte, z- oder ipsativ standardisierte Werte*) und
- welche **statistischen Größen**/graphischen Darstellungen (*Boxplots, Mittelwerte, CI, Streuung*)

schätzen Lehrkräfte bei der Rezeption von Daten aus der Unterrichtsdiagnostik als *informativ und gut interpretierbar* ein?



Vorliegende Studie

Stichprobe:

- $N_{Level_1} = 9730$ Unterrichtsbewertungen durch SuS anhand des Students Evaluation of Education Questionnaire (*SEEQ, Marsh, 1982*)
 - 5-faktorielle ordinale CFA: $\chi^2 = 3793.835$, $df = 123$, $CFI = 0.99$, $TLI = 0.987$, $RMSEA = 0.059$, $SRMR = 0.048$
- Logdaten von $N_{Level_2} = 232$ (Nachhilfe-)Mathematiklehrer*innen, Selbstauskünfte von $N_{Level_2} = 89$
 - Positivselektion Mathematikstudierender



Vorliegende Studie

- Material
 - Online Applikation (siehe auch [Demoversion unter bit.ly/Merk_GEBF18_demo](https://bit.ly/Merk_GEBF18_demo))
 - Erfassung der wahrgenommen Informativität und Interpretationssicherheit mit Single-Items

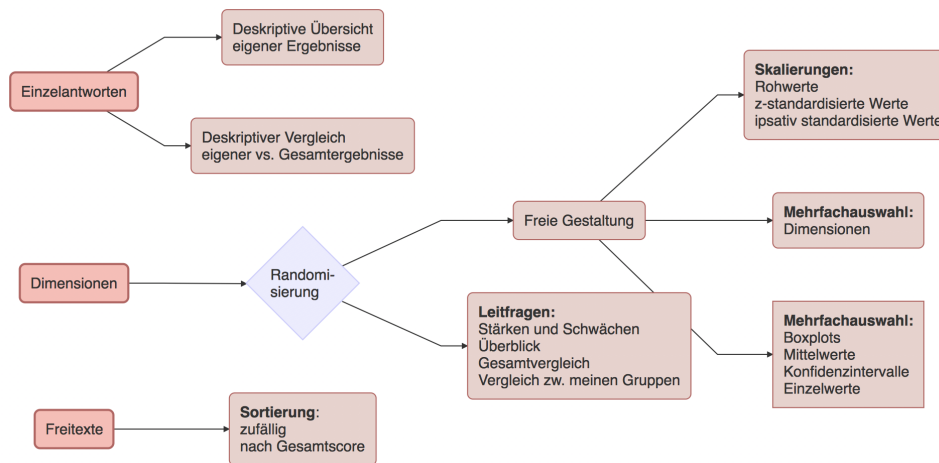


Abb. 5: Struktur der Onlineapplikation



Erste Ergebnisse

Verweildauer und Informativität

Y

Verweildauer

X

Inferenzniveau

Skalierung

Rohwerte

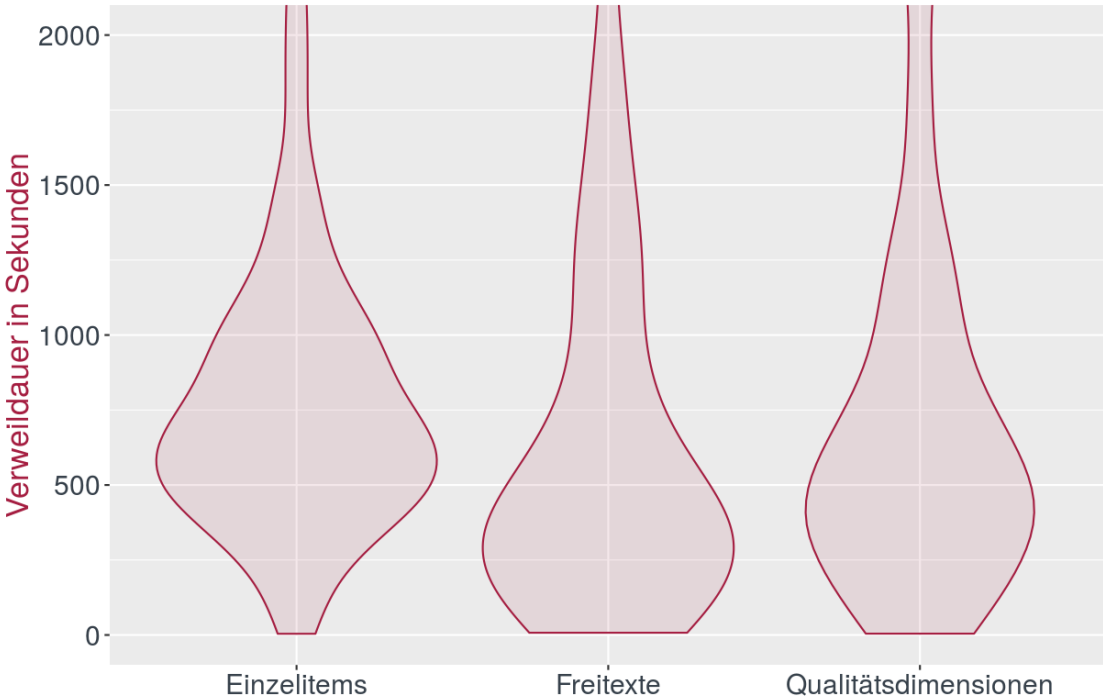
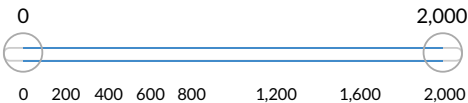
☐ Jitter

☐ MW

☐ CI

☒ Dichte

y-Limits



Limitationen

- Explorativer Charakter
- Spezifische Stichprobe
- Online-Studie



Diskussion & Ausblick

- Rezeptionsprozess als *Desiderat und Potential*
- Sind intuitiv rezipierbare *und zugleich* aussagekräftige Rückmeldeformate möglich?
 - Common language effect sizes (deskriptiv)
 - Bayes Factors (inferenzstatistisch)
- Weitere aktuelle Forschungsbemühungen
 - Experimentelle Studien zum Rückmeldezeitpunkt und zur Verbindung offener Antworten mit Globalscores
 - Replikation mit Lehramtsstudierenden
 - Effekte der (und Interventionen zur) Datenkompetenz



Vielen Dank für Ihre Aufmerksamkeit! Rückfragen?

Literatur

- Gärtner, H. (2013). Praxis und Verhältnis interner und externer Evaluation im Schulsystem im internationalen Vergleich. *Zeitschrift Für Erziehungswissenschaft*, 16(4), 693–712. online; gedruckt.
- Gärtner, H. (2016). Welche schulischen Merkmale beeinflussen die Unterrichtsqualität? Sekundäranalysen auf Grundlage von Schulinspektionsdaten. *Zeitschrift Für Erziehungswissenschaft*, 19(3), 509–526.
- Helmke, A., & Hosenfeld, I. (2005). Standardbezogene Unterrichtsevaluation. In G. Brägger, B. Bucher, & N. Landwehr (Eds.), *Schlüsselfragen zur externen schulevaluation* (pp. 127–151). Bern: h.e.p.-Verlag.
- KMK. (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Link Luchterhand.
- Koch, U. (2011). *Verstehen Lehrkräfte Rückmeldungen aus Vergleichsarbeiten? Datenkompetenz von Lehrkräften und die Nutzung von Ergebnissrückmeldungen aus Vergleichsarbeiten*. Münster [u.a.]: Waxmann.
- Maier, U. (2008). Vergleichsarbeiten im vergleich - Akzeptanz und wahrgenommener nutzen standardbasierter leistungsmessungen in Baden-Württemberg und Thüringen. *Zeitschrift Für Erziehungswissenschaft*, 11(3), 453–474.
- Maier, U. (2010). Accountability policies and teachers' acceptance and usage of school performance feedback: A comparative study. *School Effectiveness and School Improvement*, 21(2), 145–165.
- Mandinach, E. B., & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366–376.
- Marsh, H. W. (1982). Seeq: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology*, 52(1), 77–95.
- The Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher*, 32, 5–8.
- Verhaeghe, G., Vanhoof, J., Valcke, M., & Van Petegem, P. (2010). Using school performance feedback: perceptions of primary school principals. *School Effectiveness and School Improvement*, 21(July 2015), 167–188.
- Visscher, A. J. (2002). A framework for studying school performance feedback systems. In A. J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 41–71). Lisse, The Netherlands: Swets & Zeitlinger.
- Wurster, S., Richter, D., & Lenski, A. E. (2017). Datenbasierte Unterrichtsentwicklung und ihr Zusammenhang zur Schülerleistung. *Zeitschrift Für Erziehungswissenschaft*, 1–23.

