

# Zweitveröffentlichung



Helmke, Andreas; Helmke, Tuyet; Heyne, Nora; Nordheider, Iris; Schrader, Friedrich-W.

## Analyse und Bewertung von Fremdsprachenunterricht mit Hilfe von standardisierten Erhebungsinstrumenten

Datum der Zweitveröffentlichung: 23.06.2023

Verlagsversion (Version of Record), Beitrag in Sammelwerk

Persistenter Identifikator: urn:nbn:de:bvb:473-irb-598706

### Erstveröffentlichung

Helmke, Andreas; Helmke, Tuyet; Heyne, Nora; Nordheider, Iris; Schrader, Friedrich-Wilhelm: Analyse und Bewertung von Fremdsprachenunterricht mit Hilfe von standardisierten Erhebungsinstrumenten. In: Praktische Handreichung für Fremdsprachenlehrer. Jung, Udo O. H. (Hg). Frankfurt am Main : Lang, 2006. S. 537-541.

### Rechtehinweis

Dieses Werk ist durch das Urheberrecht und/oder die Angabe einer Lizenz geschützt. Es steht Ihnen frei, dieses Werk auf jede Art und Weise zu nutzen, die durch die für Sie geltende Gesetzgebung zum Urheberrecht und/oder durch die Lizenz erlaubt ist. Für andere Verwendungszwecke müssen Sie die Erlaubnis des/der Rechteinhaber(s) einholen.

Für dieses Dokument gilt das deutsche Urheberrecht.

## ANALYSE UND BEWERTUNG VON FREMSPRACHENUNTERRICHT MIT HILFE VON STANDARDISIERTEN ERHEBUNGSINSTRUMENTEN

Andreas Helmke, Tuyet Helmke, Nora Heyne  
Iris Nordheider & Friedrich-Wilhelm Schrader

*Universität Koblenz-Landau, Campus Landau*

Die Verbesserung (oder „Entwicklung“) des Unterrichts setzt eine sorgfältige, von subjektiven Verzerrungen weitgehend freie Diagnose des Ist-Standes voraus. Dafür eignen sich standardisierte Erhebungsinstrumente, bei denen man drei Ansätze unterscheidet: Die Lehrkraft holt Feedback der Schüler ein, lässt ihren Unterricht von Dritten anhand vorgegebener Kategorien einschätzen oder hinsichtlich genau eingrenzbarer Beobachtungsaspekte beschreiben. Der Artikel skizziert Grundsätze der Unterrichtsdiagnostik mithilfe standardisierter Verfahren und stellt exemplarisch einige Instrumente vor.

### 1. Die empirische Wende und ihre Folgen

Seit der empirischen Wende in der Bildungspolitik setzt sich zunehmend auch in der Schulpraxis eine an empirischen Ergebnissen orientierte Sichtweise durch. Damit geht die Erkenntnis einher, dass die Verbesserung (oder „Entwicklung“) des Unterrichts eine sorgfältige Diagnose voraussetzt. Anders als im angloamerikanischen Sprachraum ist dieser Gedanke in Deutschland noch neu. Angesichts der von der Kultusministerkonferenz (KMK) formulierten Bildungsstandards und in Anbetracht immer häufiger werdender Vergleichsarbeiten und Lernstandserhebungen kommen in den nächsten Jahren neue diagnostische Herausforderungen auf viele Lehrkräfte zu.

Engagierten Lehrpersonen sollen mit diesem Beitrag praktische Verfahren an die Hand gegeben werden. Die im Folgenden vorgestellten Verfahren des Schülerfeedbacks und der Unterrichtsbeurteilung und -beobachtung können dazu genutzt werden, den eigenen Unterricht objektiver

zu analysieren, zu bewerten und auf einer empirischen Grundlage zu verbessern. Es gab bereits in der Vergangenheit Bestrebungen, den Fremdsprachenunterricht beobachtungsbasiert zu beschreiben (Freudenstein & Pürschel 1978). Die Entwicklung standardisierter Erhebungsinstrumente und ihr Einsatz im Rahmen großer Feldstudien ist jedoch neueren Datums. In der gängigen Literatur findet man überwiegend Verfahren, die fachübergreifende Merkmale guten Unterrichts erfassen (Helmke 2006). Die nachfolgend dargestellten Verfahren thematisieren zusätzlich auch Aspekte des Fremdsprachenunterrichts.

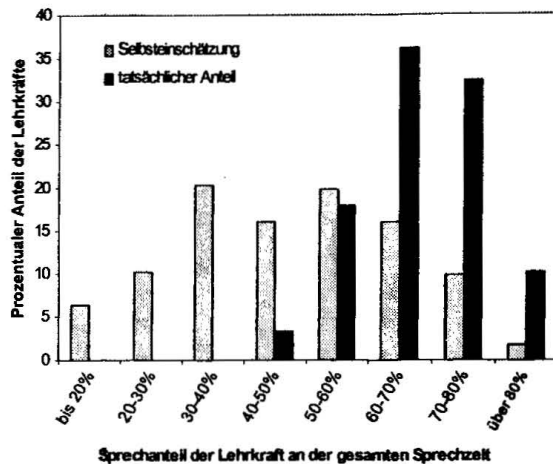
### 2. Warum standardisierte Verfahren?

Standardisierte Verfahren schalten Störquellen, die die genaue Erfassung eines Personen- oder Unterrichtsmerkmals beeinträchtigen können, so weit wie möglich aus. *Standardisiert* heißt: Nicht nur das Testverfahren (Beispiel: PISA) muss für alle untersuchten Personen exakt vergleichbar sein, sondern auch seine Durchführung (gleiche Testanweisungen, gleiches Verhalten des Testleiters). Damit soll sichergestellt werden, dass unterschiedliche Testergebnisse nicht auf unterschiedliche Messinstrumente oder Durchführungsbedingungen zurückgehen, sondern Unterschiede im zu erfassenden Merkmal widerspiegeln. Für Beobachtungsverfahren gilt entsprechendes: Einige der im Folgenden berichteten Erhebungsinstrumente wurden in großen Forschungsprojekten eingesetzt und bieten die Möglichkeit, die eigenen Ergebnisse mit dem Durchschnitt zu vergleichen, um so den eigenen (Qualitäts-)Standort zu bestimmen.

### 3. Lehrkräfte als Unterrichtsdiagnostiker

Um Unterricht zu diagnostizieren, reicht es nicht aus, seine Stärken und Schwächen aus der eigenen Erinnerung heraus einzuschätzen. Die Erinnerung ist oft trügerisch, zumal die eigene Wahrnehmung in komplexen Unterrichtssituationen stark beansprucht wird. Wie sehr Lehrpersonen sich täuschen können, wenn es um die Einschätzung des eigenen Unterrichts geht, hat kürzlich die PISA-Ergänzungsstudie DESI (Deutsch Englisch Schülerleistungen International) deutlich gemacht, bei der u.a. die Lehrereinschätzungen mit den aus Videoaufzeichnungen gewonnenen Daten zum Unterricht verglichen wurden. Die Frage, wie hoch der eigene Sprechanteil gemessen an der gesamten Sprechzeit in der vorangegangenen (videografierten) Englischstunde eingeschätzt werde, förderte bemerkenswerte Ergebnisse zutage (siehe Abbildung): Etwa ein Drittel der Lehrkräfte schätzt den eigenen Anteil an der

Sprechzeit als gering ein (höchstens 40%) – solch geringe Sprechquoten kommen jedoch gar nicht vor, wie aus dem Fehlen der schwarzen Balken im linken Teil der Abbildung ersichtlich ist. Tatsächlich liegt der Sprechanteil der Lehrpersonen durchschnittlich bei 68,3%.



Eine andere Schwierigkeit bei der Selbstdiagnose des Unterrichts ist das Fehlen eines Bezugssystems mit definierten Kriterien, anhand dessen Lehrkräfte ihren Unterricht einordnen und so mit dem von Kollegen vergleichen oder Veränderungen im eigenen Verhalten über die Zeit hinweg beurteilen können.

Anhand welcher Kriterien (Dimensionen, Merkmale) lässt sich Unterricht vergleichend beurteilen? In der *Schulpraxis* liegt der Schwerpunkt meist darauf, welche Methoden eingesetzt werden, um Unterrichtsabläufe lernwirksam zu gestalten. Diese Sichtweise ist von der Allgemeinen Didaktik geprägt: Ein Unterricht ist dann und in dem Maße gut, in dem sein Ablauf bestimmten Prinzipien und Regeln der Gestaltung entspricht, die bei der Planung und Realisierung des Unterrichts beachtet werden müssen. In der *Empirischen Unterrichtsforschung* dagegen charakterisiert man Unterricht im Hinblick auf bestimmte Qualitätsdimensionen (z.B. Strukturiertheit, Verständlichkeit, Motivierung), die nachweislich eine Rolle für den Lernerfolg spielen. Es handelt sich um eine *variablenorientierte* Betrachtungsweise, die davon ausgeht, dass sich der Unterricht verschiedener Lehrpersonen hinsichtlich bestimmter Merkmale (Variablen), die von Lehrkraft zu Lehrkraft, aber auch bei der gleichen Lehrkraft in verschiedenen Situationen (Fächer, Klassen, Zeitpunkte) variieren können (z.B. hohe Ausprägung der Variable „Verständlichkeit“ bei Lehrperson A, niedrige Ausprägung dieser Variable bei Lehr-

person B). Die empirische Unterrichtsforschung untersucht, ob unterschiedliche Ausprägungen solcher Variablen mit Unterschieden im Lernerfolg (z.B. Leistungszuwachs, Verbesserung der Lernfreude) einhergehen – eine Perspektive, die auch in der *output*-Orientierung der neuen Kernlehrpläne ihren Niederschlag findet.

Zwischen beiden Sichtweisen, der der Allgemeinen Didaktik und der Empirischen Unterrichtsforschung – Terhart (2002) charakterisiert ihr Verhältnis treffend als das zweier fremder Schwestern – liegt eine Kluft: Der Allgemeinen Didaktik ist das Denken in variablen Qualitätskategorien und erst recht die Orientierung am empirisch nachweislichen Output ebenso fremd wie der Empirischen Unterrichtsforschung das nicht empirisch basierte normative Denken der Allgemeinen Didaktik. Dazu kommt, dass das Wissen über die Wichtigkeit bestimmter Qualitätskategorien lediglich Steuerwissen ist – direkte und rezepthafte Konsequenzen für den konkreten Ablauf einer Unterrichtsstunde lassen sich daraus nicht ableiten.

#### 4. Drei Ansätze zur Analyse und Bewertung des Fremdsprachenunterrichts

Welche Möglichkeiten hat man als Lehrkraft, Aufschluss über den eigenen Unterricht zu erhalten? Im Folgenden möchten wir einige konkrete und praktisch umsetzbare Wege aufzeigen, wie der Unterricht über subjektive Urteile und narrative Beschreibungen hinaus mit standardisierten Erhebungsinstrumenten erfasst werden kann. Prinzipiell gibt es drei mögliche Ansätze: Die Lehrkraft kann ein Feedback ihrer Schüler einholen, ihren Unterricht von einem unbeteiligten Beobachter (z.B. einem Kollegen) anhand vorgegebener Kategorien einschätzen lassen (Rating) oder hinsichtlich genau eingrenzbarer Beobachtungsaspekte (z.B. „Wie oft spreche ich / sprechen die Schüler Deutsch?“) beschreiben lassen.

Welche Methode im Einzelfall in Frage kommt, hängt von der Problem- oder Interessenlage, aber auch von der Möglichkeit und Bereitschaft der Lehrkraft ab, diese in ihrem Unterricht auszuprobieren – oft führt auch ein Methoden-Mix zum Ziel. Wichtig ist nur, dass überhaupt der Schritt gemacht wird, den eigenen Unterricht systematisch zu verbessern – dazu möchten wir mit Hilfe der folgenden Verfahren ermutigen.

##### 4.1 Unterrichtswahrnehmung durch Schüler (Schülerfeedback)

Die Rückmeldung der Schüler zum eigenen Unterricht wird – z.B. mit Hilfe von Arbeitsaufträgen vom Typ „Was findet ihr gut, was könnte

ich noch besser machen?“ – bereits von vielen Kollegen eingeholt. Aussagekräftiger werden die Ergebnisse, wenn Antwortmöglichkeiten vorgegeben werden: Die ausgezählten (und evtl. sogar statistisch aufbereiteten) Ergebnisse ermöglichen Vergleiche (z.B. mit der Rückmeldung anderer Lehrkräfte, mit Rückmeldungen früherer Klassen oder mit der gleichen Klasse zu mehreren Zeitpunkten). Je nach Fragestellung können die Schülerangaben auf einen längeren Zeitraum (z.B. das gesamte zurückliegende Schuljahr) oder auf eine konkrete Unterrichtsstunde bezogen sein.

Ein Beispiel eines solchen Schülerfragebogens für den Englischunterricht stammt aus der Studie DESI. Hier exemplarisch drei Fragen mit den dazugehörigen Antwortkategorien:

*Wichtigkeit von Kommunikation:*

- Wie wichtig ist es dem/der Englischlehrer/in, dass ihr euch auf Englisch unterhalten könnt?
- Wie wichtig ist es dem/der Englischlehrer/in, dass ihr auf Englisch flüssig und ohne zu stocken sprechen könnt, auch wenn es sprachlich nicht völlig korrekt ist?

(unwichtig – eher unwichtig – eher wichtig – sehr wichtig)

*Wichtigkeit von Korrektheit:*

- Wie wichtig ist es dem/r Englischlehrer/in, dass eure Aussprache korrekt ist?
- Wie wichtig ist es dem/r Englischlehrer/in, dass eure Sätze beim Sprechen grammatikalisch korrekt sind?

(unwichtig – eher unwichtig – eher wichtig – sehr wichtig)

*Instrumentelle Motivierung:*

- Mein/e Englischlehrer/in weist darauf hin, wie nützlich Englisch im Alltag ist.
- Mein/e Englischlehrer/in betont, dass Englisch in vielen Berufen eine große Rolle spielt.

(stimmt gar nicht – stimmt eher nicht – stimmt eher – stimmt ganz genau)

Schüler sind zwar keine Experten für didaktische Fragen, können aber viele Aspekte des Unterrichts wie Verständlichkeit oder Unterstützung gut beurteilen. Da es sich beim Schülerfeedback nicht um Unterrichtswahrnehmungen einzelner Schüler, sondern um die gemittelten Angaben der gesamten Klasse handelt, und den Schülerurteilen oft auch längerfristige Beobachtungszeiträume zugrunde liegen, sind die Angaben meistens recht zuverlässig und aussagekräftig. Genauere Hinweise zu den Grundlagen und den prakti-

sehen Einsatzmöglichkeiten von Schülerbeurteilungen finden sich bei Helmke (2004). Aus dem Internet, insbesondere von den Bildungsservern der Länder bzw. Kantone können viele Instrumente für Schülerfeedback herunter geladen und genutzt werden (URL 2 und 3).

#### 4.2 Kategorienbasierte Beurteilung des Unterrichts

Eine andere Möglichkeit besteht darin, dass ein unbeteiligter, aber qualifizierter Beobachter den Unterricht bezüglich allgemeiner Qualitätsmerkmale (Rating) beurteilt. Dazu kann man z.B. einen Kollegen bitten, in einer Stunde zu hospitieren und anschließend den Unterricht anhand eines Ratingbogens hinsichtlich bestimmter vorher festgelegter Aspekte zu beurteilen. Eine solche kategorienbasierte Evaluation hat den Vorteil, dass die Aufmerksamkeit des Beobachters konsequent auf bestimmte Qualitätsdimensionen gelenkt wird, z.B. solche, die für den Lernerfolg der Schüler nachweislich entscheidend sind. Kennzeichen dieser Verfahren ist ein relativ hoher Beurteilungsspielraum, da vom konkreten Unterrichtsgeschehen auf bestimmte Kategorien geschlossen werden muss, was vom Beobachter ein bestimmtes Maß an Sachkunde und eine hinreichende Vertrautheit mit den verwendeten Kategorien erfordert. Ratingbogen gehören bei der externen Schulevaluation (durch Inspektoren oder Evaluationsteams) inzwischen zum Standardrepertoire.

Als besonders fruchtbar hat sich ein Vorgehen herausgestellt, bei dem die Ratings eines Fremdbeobachters (z.B. eines hospitierenden Kollegen) mit der Selbsteinschätzung der unterrichtenden Lehrperson anhand des gleichen Ratingbogens verglichen werden. Differenzen zwischen den beiden Rating-Profilen können dann Anlass für fachliche, fachdidaktische oder pädagogische Diskussionen sein. Vorgestellt werden soll hier exemplarisch der Bereich „Zielorientierung“ mit den untergeordneten Unterrichtsmerkmalen

- Kommunikation (Schaffung von Gesprächssituationen im Unterricht)
- Anregung von Lern- und Denkstrategien
- Eingehen auf Lern- und Verständnisschwierigkeiten
- Anregung anspruchsvoller, weiterführender Themen
- Berücksichtigung von Flüssigkeit
- Berücksichtigung von Korrektheit

Zu allen zu beurteilenden Merkmalen werden im DESI-Ratingbogen kurze Definitionen sowie Antwortkategorien vorgegeben. So wird beispiels-

weise die *Berücksichtigung von Korrektheit* wie folgt definiert:

„Die Lehrkraft gibt zu erkennen, dass ihr Korrektheit/Präzision in der Spracharbeitsphase wichtig ist; sie achtet auf die Korrektheit/Präzision der Äußerungen der Schüler; in Sprachlernphasen macht sie deutlich, wenn Fehler vorkommen, korrigiert diese oder lässt sie korrigieren“.

Die Einschätzung dazu erfolgt anhand der folgenden Antwortoptionen:

- trifft zu: es wird auf nahezu alle Fehler eingegangen; bei Auftreten von Fehlern geht die Lehrkraft selbst darauf ein oder lässt Schüler darauf eingehen
- trifft eher zu: es wird häufig auf Fehler eingegangen
- trifft eher nicht zu: es wird nur gelegentlich auf Fehler eingegangen
- trifft nicht zu: es wird so gut wie gar nicht auf Fehler eingegangen; die Lehrkraft beachtet Fehler nicht weiter; korrigiert kaum
- Beurteilung nicht möglich, da keine sprachlernorientierte Phase

Das komplette Ratinginstrument wird im Laufe des Jahres 2006 in einer vom DIPF herausgegebenen Reihe erscheinen, in der alle Untersuchungsinstrumente dokumentiert sind (URL 4).

#### 4.3 Beobachtung des Unterrichts

Eine dritte Herangehensweise zur Analyse und Bewertung von Unterricht bieten Beobachtungsverfahren. Ein im angloamerikanischen Raum verbreitetes Instrument, das speziell für den Fremdsprachenunterricht entwickelt wurde, ist COLT (*Communicative Orientation of Language Teaching Observation Scheme*) von Spada & Fröhlich (1995). Dieses Verfahren stellt zwei Beobachtungspläne bereit, mit denen Beobachter durch Ankreuzen vorgegebener Kategorien den Unterricht beschreiben. Zudem beinhaltet es Hinweise zur Auswertung dieser Aufzeichnungen. So sieht Beobachtungsbogen A vor, dass die in Zeitintervalle unterteilte Unterrichtsstunde abschnittsweise anhand folgender Kategorien kodiert wird:

- Teilnehmer (Lehrer, Schüler, Lehrer-Schüler, Schüler-Schüler, gesamte Klasse)
- Aufgabenstellung (für alle gleich, verschiedenen)
- Gegenstand (Fokus der Sprache: Form, Funktion, Diskurs, Soziolinguistik)
- Themenbestimmung (durch Lehrer/Text, durch Lehrer-Schüler gemeinsam, durch Schüler)

- Tätigkeit der Schüler (zuhören, sprechen, lesen, schreiben)
- Materialien (Typ: Text, Ton, Bild; Quelle: didaktisch entwickelt, original in Muttersprache, von Schülern entwickelt)

Mithilfe des Beobachtungsbogens B werden Merkmale der verbalen Lehrer-Schüler-Interaktion beschrieben:

- Unterrichtssprache
- Informationsgehalt (vorhersagbar/unvorhersagbar, pseudo/echt)
- Länge der Äußerung (Ein-Wort-Sätze/Satzfragment usw.)
- Art der Äußerung (Korrektur, Wiederholung, Umschreibung, Kommentar, Erweiterung, Nachfrage zur Klärung, Nachfrage für weitere Informationen)

Das Instrument COLT war Grundlage des Unterrichtsbeobachtungsinventars, das für das Projekt DESI weiterentwickelt und differenziert wurde. Bei der sogenannten Basiskodierung des Englischunterrichts werden grundlegende Merkmale des Unterrichtsgeschehens präzise (d.h. Zeitannteil und Vorkommenshäufigkeit) erfasst, wie zum Beispiel:

- Art von Schüleräußerungen (frei sprechen, nachsprechen, wiederholen, ablesen)
- Selbstkorrektur von Fehlern durch Schülerinnen und Schüler
- Umgang mit Schülerfehlern
- Gehalt der Lehrer-Rückmeldung
- Lehrer-Schüler-Dialoge, die sich über mehrere Stationen erstrecken

Deskriptive Ergebnisse dieser und anderer Merkmale des Englischunterrichts finden sich unter Zentrale Befunde der DESI-Studie im Internet (URL 1).

#### 5. Weiterführende Literatur

Das Lehrbuch von Helmke (2004) richtet sich an Lehrkräfte und bietet einen Überblick zum Thema Unterrichtsqualität. Vertiefungsmöglichkeiten im Bereich der Unterrichtsbeobachtung finden sich bei Waxman, Tharp & Hilberg (2004). Einen praxisnahen Überblick über Evaluationsmethoden im Unterrichtsalltag verschaffen Burkard & Eikenbusch (2000). Mit Fragen der Beobachtung des Fremdsprachenunterrichts beschäftigt sich Ziebell (2003).

#### 6. Ausblick

Die DESI-Studie des Englischunterrichts hat deutlich gemacht, wie wenig sich Lehrpersonen ihres

eigenen Unterrichts bewusst sind und wie selten kooperiert wird. So geben z.B. lediglich 2% der Englischlehrkräfte an, gelegentlich (d.h. mindestens mehrmals pro Monat) in kollegialer Zusammenarbeit ihren Unterricht vorzubereiten. Der Einsatz standardisierter Verfahren zur Beschreibung und Bewertung des Unterrichts ist nicht nur ein aussichtsreicher Weg, um die Qualität des Unterrichts zu erfassen und zu steigern, sondern auch ein sinnvoller Ansatz und Anlass für schulinterne Kooperationen. Darauf basierende Projekte lassen sich ausgezeichnet in Schulprogramme integrieren, die einen Schwerpunkt auf das Thema „Unterricht“ legen.

Besonders ergiebig ist der Einsatz solcher Verfahren auf der Basis von Videoaufzeichnungen, die inzwischen ausgesprochen unkompliziert und preiswert zu erstellen sind. Vor allem bieten sie die Möglichkeit, den Unterricht detailliert, wiederholt, im Hinblick auf verschiedene Aspekte, unter verschiedenen Perspektiven und mit unterschiedlichen Methoden zu analysieren (Helmke & Helmke 2004). Die Unterrichtsvideografie stellt daher heute – in der Unterrichtsforschung wie in der Unterrichtspraxis – zweifellos den Königsweg zur Evaluation und Diagnose des Unterrichts dar und könnte ein wichtiger Motor für die Professionalisierung der Lehrkräfte sein.

#### LITERATURVERZEICHNIS

- Burkard, C. & Eikenbusch, G. (2000). *Praxis- handbuch Evaluation in der Schule*. Berlin, Cornelsen.
- Freudenstein, R. & Pürschel, H. (1978). Die fremdsprachliche Unterrichtsstunde. Ein Vorschlag zur Analyse. *Praxis des neu- sprachlichen Unterrichts* 25 (3), 227-235.
- Helmke, A. (2006). Was wissen wir über guten Unterricht? *Pädagogik* 58 (2), 42-45.
- Helmke, A. (2004). *Unterrichtsqualität – erfassen, bewerten, verbessern*. Seelze, Kallmeyer-Verlag.
- Helmke, A. & Helmke, T. (2004). Videobasierte Unterrichtsreflexion. *Seminar* 10 (4), 48-66.
- Spada, N. & Fröhlich, M. (1995). *The communicative orientation of language teaching (COLT) observation scheme: coding conventions and applications*. Sydney, National Centre for English Language Teaching and Research.
- Terhart, E. (2002). Fremde Schwestern: Zum Verhältnis von Allgemeiner Didaktik und empirischer Lehr-Lern-Forschung. *Zeitschrift für Pädagogische Psychologie* 16 (2), 77-86.
- Waxman, H.C. et al. (eds.) (2004). *Observational research in U.S. classrooms. New approaches for understanding cultural and linguistic diversity*. Cambridge, Cambridge University Press.

Ziebell, B. (2003). *Unterrichtsbeobachtung und Lehrerverhalten (mit Videosequenzen)*. Berlin, Langenscheidt.

#### URLs

- 1 <http://www.dipf.de/desi>
- 2 <http://www.qis.at/sitemap2.asp>
- 3 <http://www.das-macht-schule.de/seis-instrument/erhebungsinstrumente/fragebogen>
- 4 <http://www.dipf.de>