

# QUALITÄTSSICHERUNG IN DER BEGABUNGSFÖRDERUNG PILOTSTUDIE ZUR ENTWICKLUNG EINES QUALITÄTSANALYSE- INSTRUMENTARIUMS FÜR DEN PFLICHTSCHULBEREICH

Veröffentlicht in: news&science. Begabtenförderung und Begabungsforschung. özbf, Nr. 22/Ausgabe 2, 2009, S. 27-33.

## Einleitung

Überlegungen zur Qualitätsentwicklung in der Begabungsförderung sind nicht von Fragen zur Unterrichtsqualität im Allgemeinen zu trennen. Wissenschaftstheoretisch wird vor allem im Rahmen hochschuldidaktischer Bemühungen die Frage nach der „guten“ Lehre aufgegriffen (Berendt 1994; Webler 1991).

Qualitätssicherung, die dem klassischen „Input – Prozess – Output“-Paradigma unterliegt, bedarf der Evaluation, die sich mit der Wirkung gesetzter Maßnahmen beschäftigt. (vgl. Wottawa/Thierau 1990, S. 9)

Nur selten wird bei Überlegungen dieser Art allerdings berücksichtigt, wie es überhaupt um die Güte des eingesetzten Messinstrumentes und damit auch um die Gültigkeit der Messergebnisse bestellt ist. Soll nämlich wie im Folgenden das studentische Urteil/Schüler/innenurteil tatsächlich Aufschluss über die „Qualität“ der Lehre erbringen, müssen sich die Studierenden oder Schüler/innen zum einen darüber einig sein, was „gute“ Lehre ist (Validität des Urteils), zum anderen müssen die Kriterien der Studierenden bzw. Schüler/innen einen objektiven Maßstab (Objektivität des Urteils) für „gute“ Lehre darstellen (Schweer 2001, S. 160).

Im folgenden Beitrag soll ein für die Praxis relevantes Qualitätsanalyse-Instrumentarium zur Diskussion gestellt werden, welches in das Bereichsfunktions-Prozessmodell (Urban 2004) eingebettet ist. Dieses versteht sich als Alternative zum TQM (Total Quality Management) bzw. EFQM (European Foundation for Quality Management) und als Gesamtsystem zur Qualitätsanalyse von Prozessen innerhalb von Bildungseinrichtungen. Entwickelt für den Hochschulbereich, soll das Qualitätsanalyse-Instrumentarium für die Pflichtschule adaptiert werden und Möglichkeiten einer Evaluation von Qualitätskriterien in der Begabungsförderung vorgestellt werden.

## Qualitätsbegriff

Die vorliegende Pilotstudie greift auf einen prozessorientierten Qualitätsbegriff zurück, wie er im Bereichsfunktions-Prozessmodell von Urban (2004, S. 59-89) anschaulich dargestellt ist (siehe Abb. 1). Im Mittelpunkt steht die Qualität der Lehre/Unterrichtsqualität, die aber nur einen Teil des gesamten Evaluationsprozesses darstellt, der selbst wieder von anderen Bereichen prozesshaft beeinflusst wird und seinerseits diverse Wechselwirkungen auf andere Bereiche – vor allem auf die Ergebnisse – ausübt.

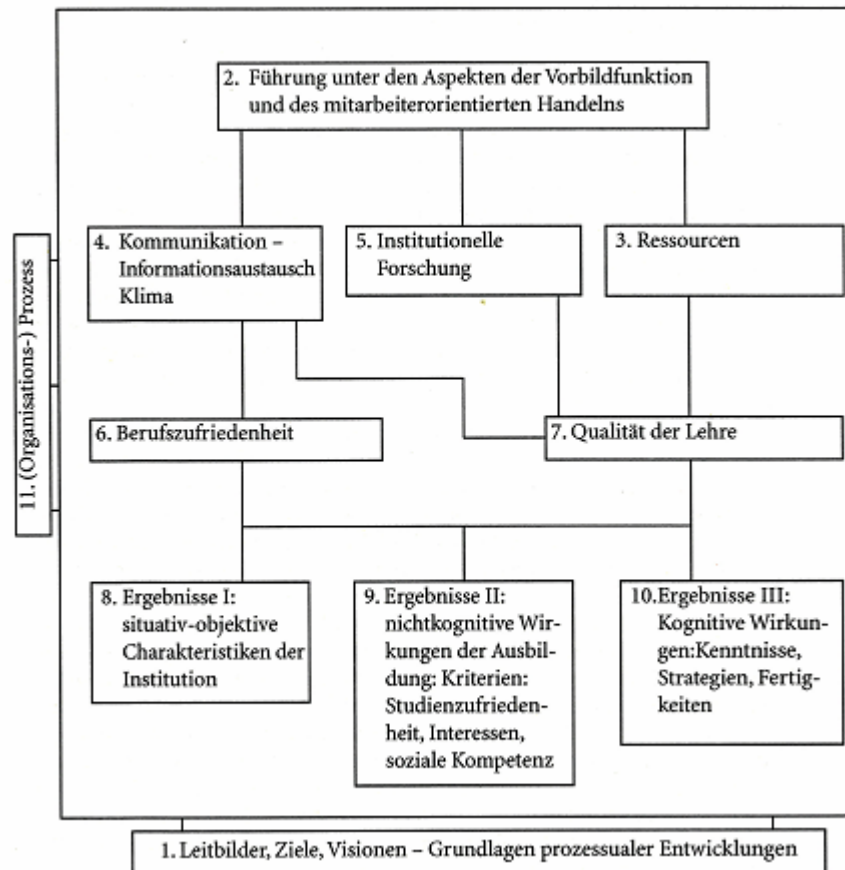


Abb. 1: Bereichsfunktions-Prozessmodell

Zur Definition von Unterrichtsqualität eignet sich das dargestellte Modell jedoch nicht, da es nicht sämtliche Zwischenstationen bezüglich des Zustandekommens von Studierenden- bzw. Schüler/innenleistungen erfasst. Dazu bedient sich Urban (2004, S. 123-126) des „Angebots-Nutzungs-Modells“ von Helmke (1996, zit. nach Helmke 2009), das den Komplex „Lehre-Ergebnis“ besser definiert. Der Schwerpunkt liegt auf der formativen Evaluation<sup>1</sup>, d. h. Evaluation einzelner Lehrveranstaltungen, mit dem Ziel, den Lehr- und Lernprozess in ständigem Austausch mit den Studierenden/Schülerinnen und Schülern zu optimieren. In diesem Modell stehen Prozessmerkmalen (guter, effizienter Unterricht, Lehrperson, Lehrangebot) Produktmerkmalen (realer und subjektiv eingeschätzter Lehrerfolg, Studienzufriedenheit) gegenüber, dazwischen liegen Vermittlungsprozesse. Ob und in welcher Weise man von den Lehrveranstaltungen profitiert, hängt neben motivationalen Eingangsvoraussetzungen davon ab,

- warum man sich für eine Lehrveranstaltung entschieden hat,
- wie regelmäßig man teilnimmt,
- wie intensiv die Lehrveranstaltung genutzt wird, d. h. wie intensiv das Lernengagement ist,
- wie sehr die Passung zwischen Person des Lehrenden und den Methoden einerseits und den individuellen Eingangsvoraussetzungen der Studierenden/Lernenden andererseits gegeben ist. Lernerfolg und Zufriedenheit hängt somit auch von den individuellen Lernstilen und Präferenzen ab.

Die spezielle Ausbildungssituation an den ehemaligen Pädagogischen Akademien – hier führte Urban (2004) seine Untersuchung durch – lässt ihn auf die auf Freiwilligkeit basierenden Vermittlungsprozesse weitgehend verzichten. Das von ihm entwickelte Instru-

<sup>1</sup> Dem gegenüber steht eine summative Evaluation nach Abschluss eines Projekts oder eines Innovationsvorhabens etc. (Altrichter et al. 2006, S. 24).

mentarium zur studentischen Lehrevaluation (a. a. O., S. 267f.) gliedert sich in 3 Bereiche:

- 1) Prozessmerkmale im Unterricht  
Skala 1 enthält Items, die die didaktische Kompetenz der Lehrer/innen aus Sicht der Studierenden abtestet (Reliabilitätskoeffizient = .83).  
Skala 2 enthält Items, die das Engagement und die Teilnehmerorientiertheit der Lehrer/innen aus Sicht der Studierenden abtestet (Reliabilitätskoeffizient = .87).  
Skala 3 enthält Items, die die Fachkompetenz der Lehrenden aus Sicht der Studierenden abtestet (Reliabilitätskoeffizient = .76).
- 2) Produktmerkmale im Motivations- und Interessensbereich  
Skala 4 enthält Items, die die Motivierung der Studierenden abtestet (Reliabilitätskoeffizient = .74).
- 3) Produktmerkmal Zufriedenheit  
Skala 5 enthält Items, die die Zufriedenheit – als Zufriedenheitsmaß dient hier die Zufriedenheit mit der individuellen Kompetenzvermittlung – abtestet (noch keine Messwerte für die Reliabilität).

Obiges Evaluationsinstrumentarium steht in Abhängigkeit der Validität studentischer Lehrevaluation. Untersuchungen Urbans (2004) zum Einfluss der Persönlichkeitsvariablen zeigen auf, dass die Beurteilung der Ausbildungssituation von der erlebten Selbstverwirklichung und vom Selbstkonzept der eigenen Leistung wesentlich beeinflusst wird. Clusteranalysen machen deutlich, dass Studierende mit einem hohen Leistungskonzept und erlebter Selbstverwirklichung sowie emotionaler Stabilität und einem hohen Grad an Gewissenhaftigkeit fast alle Aspekte einer qualitativen Lehre und ihre Ergebnisse günstiger beurteilen als Studierende mit niedrigen Werten in den Persönlichkeitsvariablen. Kovarianzanalysen zeigen jedoch, dass die Mittelwerte der Qualitätsvariablen durch die Einbeziehung der Persönlichkeitsvariablen korrigierbar sind. Nach Urban (2004, S. 249) dürfte trotz aller Einwände die Beurteilung der Lehre durch Studierende auf der Basis der Mittelwerte eine valide bzw. erwartungstreue Schätzung zur Qualität abgeben.

### **Entwicklung des Schüler/innenfragebogens**

Da der Fragebogen zur studentischen Lehrevaluation inhaltlich äußerst komplex ist, wurde begonnen, die Items dem Verständnisniveau von Pflichtschüler/innen anzupassen. So wurde etwa: „Der Vortragende vermittelt den Stoff anschaulich und verständlich“ in „Der Lehrer/die Lehrerin verwendet beim Erklären gute Beispiele“ umformuliert oder „Die Lehrveranstaltungen motivieren mich, fachliche Diskussionen mit Kollegen zu führen“ in „In der Pause unterhalte ich mich mit meinen Mitschülern über Aufgaben aus dem Unterricht“. Teilweise sind die Items so komplex formuliert („Durch intensive geistige Auseinandersetzung mit relevanten Inhalten wird die Fähigkeit zu unabhängigem Denken gefördert“), dass neue Items formuliert wurden, um wie hier die Motivation der Schüler/innen abzutesten („Der Lehrer/die Lehrerin kann mich ermutigen, schwierige Aufgaben zu lösen“). Skala 3 „Professorenkompetenz“ wurde weggelassen, da nach Meinung der Verfasserin Schüler/innen im Pflichtschulbereich noch nicht in der Lage sind, die fachliche Kompetenz von Lehrenden zu beurteilen. Anders verhält es sich mit dem Faktor „Didaktische Kompetenz der Lehrer/innen“, der jedoch den Rahmenbedingungen im Pflichtschulbereich angepasst werden musste. Dementsprechend wurde der Faktor um den Kompetenzbereich Individualisierung/Differenzierung erweitert. In Anlehnung an Gruehn (2000) wurden folgende Items formuliert: „Wir Schüler bekommen vom Lehrer/von der Lehrerin unterschiedlich schwierige Aufgaben gestellt“, „Gute Schüler bekommen vom Lehrer/von der Lehrerin andere Aufgaben“, „Der Lehrer/die Lehrerin gibt mir schwierigere/leichtere Aufgaben als den übrigen Schülern“ und „Ich darf im Unterricht wählen, welche Aufgaben ich lösen möchte“.

Der Fragebogen von Urban (2004) sieht eine fünfstufige Antwortskala vor. Eine Skala mit fünf Antwortmöglichkeiten gestattet, präziser zu messen, jedoch sind nach Meinung der Verfasserin die Probandinnen und Probanden einer 4. Klasse Volksschule noch nicht in der Lage, so feine Abstimmungen vorzunehmen. Daher wurde die Antwortskala für den Schüler/innenfragebogen auf 4 Stufen gekürzt. Zudem wurde das am Ludwig-Boltzmann Institut für Schulentwicklung und international-vergleichende Schulforschung entwickelte Layout (siehe Tab.1) verwendet, nachdem die Schüler/innen aufgefordert werden, zuerst „ja“ oder „nein“ zu wählen und dann erst genauer nach „immer“, „oft“, „selten“ oder „nie“ zu unterscheiden (Hanisch 2004, S. 898).

	JA		NEIN	
	immer	oft	selten	nie
1. Die Lehrerin ermutigt mich, im Unterricht Fragen zu stellen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Die Lehrerin erklärt im Unterricht gut.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Abb. 3 : Layout des Schülerfragebogens

### Pilotstudie

Der Fragebogen wurde im Dezember 2008 an 98 Volksschüler/innen der 4. Schulstufe vorgetestet. An der Studie nahmen 46 Buben und 52 Mädchen teil. Der Anteil von Schüler/innen mit nichtdeutscher Umgangssprache lag bei 33,3 % und damit laut Statistik Austria (in <http://www.statistik.at>, 14.12.2008) unter dem Wiener Anteil von 53 %. Es muss angenommen werden, dass Eltern von Schüler/innen mit nichtdeutscher Umgangssprache aufgrund von Verständnisschwierigkeiten in einigen Fällen die Einwilligung zur Teilnahme ihrer Kinder an der Pilotstudie nicht erteilten, was zur Verzerrung der Stichprobe beigetragen haben kann. Die Stichprobe war insofern zufällig gewählt, als die Schulen nur nach dem Kriterium der zu erwartenden Bereitschaft der Schulleiter/innen<sup>2</sup> zur Teilnahme an der Pilotstudie kontaktiert wurden. Der Stichprobenumfang liegt mit  $n=98$  deutlich über dem zentralen Grenzwerttheorem von  $n \geq 30$  (vgl. Bortz 1999, S. 102).

Nach Genehmigung der Pilotstudie seitens des Stadtschulrats für Wien wurden die teilnehmenden Schulen von der Verfasserin aufgesucht, die die Untersuchung in den Klassen leitete, um etwaige Missverständnisse bei der Beantwortung der Fragen im Vorhinein ausschließen zu können. Aufgabe der qualitativen sowie quantitativen Vortestung war es, nicht brauchbare Fragen zu streichen bzw. umzuformulieren und die Güte des Fragebogens zu bewerten. Dabei sollten die befragten Schüler/innen besonders auf die Formulierung der Fragen, deren Verständlichkeit und Sinnhaftigkeit achten. Alle Einwände der Befragten wurden in der Endform des Fragebogens berücksichtigt.

### Ergebnisse

Die Datenauswertung erfolgte mittels Statistikprogramm SPSS 15.01. Vorab wurden mittels Kreuztabelle jene Schüler/innen ermittelt, die bei beiden Kontrollitems ein sozial erwünschtes Antwortverhalten gezeigt hatten. Jene 11 Schüler/innen wurden aus den weiteren Berechnungen ausgeschlossen, sodass schlussendlich die Datensätze von 87 Schüler/innen übrig blieben.

<sup>2</sup> Die Verfasserin bedankt sich bei den Schulleiterinnen/Schulleitern und Klassenlehrerinnen/Klassenlehrern der Volksschulen Pfeilgasse 42b, 1080 Wien; Märzstraße 178-180, 1140 Wien; Odoakergasse 48, 1160 Wien und Spielmannsgasse 1, 1200 Wien für die Mitwirkung an der Pilotstudie sowie die Unterstützung bei der Datenerhebung.

Die Faktorenanalyse für die Schüler/innendaten zeigte die erwartete 4-Faktorenlösung. Bei der inhaltlichen Interpretation der Faktoren konnten die Bezeichnungen „Motivation der Schüler/innen und Zufriedenheit mit der allgemeinen didaktischen Kompetenz von Lehrer/innen“, „Wahrgenommene Schülerorientiertheit und Engagement der Lehrer/innen“, „Erlebte didaktische Kompetenz der Lehrer/innen in der Individualisierung/Differenzierung“ und „Zufriedenheit der Schüler/innen mit den Lerninhalten und der eigenen Arbeitshaltung“ gefunden werden. In weiterer Folge wurden jene Faktoren einer Reliabilitätsanalyse unterzogen und das Cronbach  $\alpha$  ermittelt.

Die nach Prozess- und Produktvariablen geordneten Skalen (Faktoren) des Schüler/innenfragebogens zur Unterrichtsqualität sehen in der Endform demnach wie folgt aus:

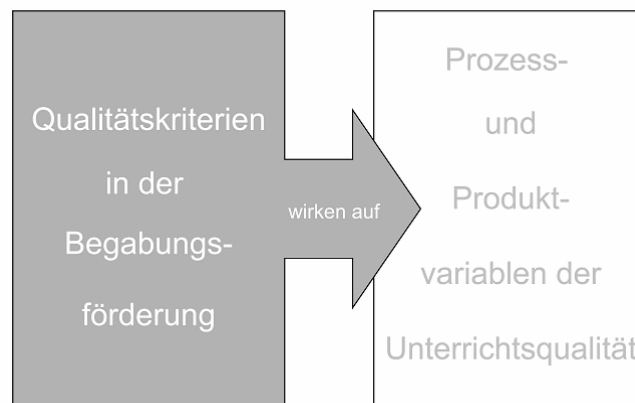


Abb. 2: Evaluationsdesign

### Prozessvariablen

Skala (Faktor) 1: „Wahrgenommene Schülerorientiertheit und Engagement der Lehrer/innen“

1. In der Pause unterhalte ich mich mit den anderen Kindern über Aufgaben aus dem Unterricht.
  2. Wenn ich etwas nicht kann, hilft mir der Lehrer/die Lehrerin.
  3. Der Lehrer/die Lehrerin verwendet beim Erklären gute Beispiele.
  4. Zu Hause beschäftige ich mich mit Arbeiten aus dem Unterricht alleine weiter.
  5. Ich darf im Unterricht wählen, welche Aufgaben ich lösen möchte.
  6. Ich darf im Unterricht Aufgaben vorschlagen, die mich interessieren.
  7. Der Lehrer/die Lehrerin interessiert sich für meine Meinung.
  8. Der Lehrer/ die Lehrerin nimmt sich für mich Zeit, wenn ich Fragen habe.
  9. Ich weiß, warum wir im Unterricht bestimmte Dinge lernen.
  10. Ich darf mein Wissen den anderen Kindern mitteilen.
  11. Der Lehrer/die Lehrerin ermutigt mich, im Unterricht Fragen zu stellen.
- (Reliabilitätskoeffizient:  $r = .777$ )

Skala (Faktor) 2: „Erlebte didaktische Kompetenz der Lehrer/innen in der Individualisierung/Differenzierung“

1. Der Lehrer/die Lehrerin gibt mir schwierigere Aufgaben als den übrigen Schülerinnen und Schülern.
2. Im Unterricht arbeite ich abwechselnd alleine, mit einem Partner oder gemeinsam mit allen Kindern. (statt: Der Lehrer/ die Lehrerin gestaltet den Unterricht abwechslungsreich.)
3. Wir Schüler/innen bekommen vom Lehrer/von der Lehrerin unterschiedlich schwierige Aufgaben.

4. Gute Schüler/innen bekommen vom Lehrer/ von der Lehrerin andere Aufgaben.
  5. Wenn ich mich im Unterricht anstrenge, kann ich die Aufgaben alleine lösen.
  6. Ich weiß, warum wir im Unterricht bestimmte Dinge lernen.
  7. Ich darf mein Wissen den anderen Kindern mitteilen.
- (Reliabilitätskoeffizient:  $r = .650$ )

### Produktvariablen

Skala (Faktor) 3: „Motivation der Schüler/innen und Zufriedenheit mit der allgemeinen didaktischen Kompetenz der Lehrer/innen“

1. Der Lehrer/die Lehrerin erklärt im Unterricht gut.
  2. Der Unterricht macht mich neugierig auf neue Aufgaben.
  3. Bei dem Lehrer/der Lehrerin lerne ich viel.
  4. Die Aufgaben im Unterricht interessieren mich.
  5. Ich freue mich auf den Unterricht.
  6. Der Unterricht macht mir Spaß.
  7. Nach dem Unterricht habe ich das Gefühl, dass ich jetzt mehr weiß.
  8. Der Lehrer/die Lehrerin erzählt spannend.
- (Reliabilitätskoeffizient:  $r = .832$ )

Skala (Faktor) 4: „Zufriedenheit der Schüler/innen mit den Lerninhalten und der eigenen Arbeitshaltung“

1. Wir Schüler/innen bekommen im Unterricht unterschiedlich schwierige Arbeitsblätter.
  2. Die Aufgaben aus dem Unterricht werde ich später noch brauchen.
  3. Ich bemühe mich, meine Arbeit ordentlich zu erledigen.
  4. Ich darf im Unterricht wählen, welche Aufgaben ich lösen möchte.
  5. Nach dem Unterricht habe ich das Gefühl, dass ich jetzt mehr weiß.
  6. Es macht mir Spaß, meine Arbeiten in der Klasse vorzustellen.
- (Reliabilitätskoeffizient:  $r = .655$ )

### Formative Evaluation

Um im Rahmen einer formativen Evaluation einzelne begabungsfördernde Angebote zu messen, kann der Begriff „Unterricht“ in den einzelnen Items des Fragebogens durch den Begriff „Kurs“ ersetzt werden. Bei Einsatz des Fragebogens im Sekundarbereich wird es notwendig sein, den Begriff „Unterricht“ für die einzelnen Fachbereiche zu präzisieren (z. B. Mathematikunterricht usw.)

Nach Urban (2004, S. 263) sind die Reliabilitätswerte von der Größe der Stichprobe weitgehend unabhängig. Demnach gelten die für die Skalen des Schüler/innenfragebogens angegebenen Reliabilitätskoeffizienten auch bei einem Einsatz in kleineren Gruppen z. B. Klassen- oder Kursgröße.

Bei deskriptiven Auswertungsmethoden (Häufigkeitsauszählung, Mittelwertsberechnung) ist es jedoch notwendig, die Items „Ich weiß, warum wir im Unterricht bestimmte Dinge lernen“ und „Ich darf mein Wissen den anderen Kindern mitteilen“ nur der Skala „Differenzierung“ zuzuordnen bzw. „Ich darf im Unterricht wählen, welche Aufgaben ich lösen möchte“ nur der Skala „Zufriedenheit der Schüler/innen“, was die Reliabilität für die Skala „Schülerorientiertheit“ auf  $r = .741$  senkt. Ebenso wird das Item „Nach dem Unterricht habe ich das Gefühl, dass ich jetzt mehr weiß“ nur der Skala „Zufriedenheit der Schüler/innen“ zugeordnet, was die Reliabilität für die Skala „Motivation der Schüler/innen“ auf  $r = .812$  senkt.

Schon bei der Entwicklung des Fragebogens wurde darauf geachtet, dass alle Items in die selbe Richtung gepolt sind. Den Lehrkräften ist es daher auch ohne eine computerunterstützte Auswertung möglich, eine Schüler/innenbewertung der einzelnen Faktoren zu ermitteln und eine sequentielle Lehr-Lern Analyse (Schweer 1996) vorzunehmen.

### Summative Evaluation<sup>3</sup>

Mit der Evaluation von Qualitätskriterien in der Begabungsförderung stößt man auf das schwierige Problem der Wirkungsforschung. Dabei gilt es jene Größen zu isolieren, die tatsächlich mit Veränderungen zusammenhängen. Nach Urban (2004, S. 48) ist ein gut geplantes multivariates Untersuchungsdesign mit den damit zusammenhängenden Auswertungsmethoden wahrscheinlich die zielführendste Maßnahme.

Das „Input – Prozess – Output“-Paradigma der quantitativen Evaluationsforschung impliziert eine gerichtete Unterschiedshypothese. Nach Helmke/Schrader (2001, S. 86) sind unabhängige Variablen zu definieren, die linear auf die abhängigen Variablen wirken.

Im konkreten Fall liegt eine positiv gerichtete Unterschiedshypothese vor, die davon ausgeht, dass vordefinierte begabungsfördernde Maßnahmen (unabhängige Variablen) positiv auf die Unterrichtsqualität (definiert durch die abhängigen Prozess- und Produktvariablen) wirken. Unter Bezugnahme auf das Bereichsfunktions-Prozessmodell sind v. a. die Variablen Ressourcen, Forschung und Kommunikation als unabhängig zu definieren. Übersetzt auf den Schulbereich können personelle Ressourcen (Anzahl der Lehrer/innen mit begabungsfördernder Ausbildung, Anzahl der Teamlehrer/innenstunden etc.), räumliche Ressourcen (Lernwerkstatt, Bibliothek etc.), begabungsfördernde Weiterbildung der Lehrer/innen (Teilnahme an Kongressen etc.), Kooperation mit internen und externen Expertinnen und Experten, Lernumgebung (Material zur Individualisierung etc.), Formen äußerer Differenzierung<sup>4</sup> (Pull-out-Programm, Drehtürmodell etc.) usw. als unabhängige Variablen definiert werden (vgl. özbf 2006; vgl. auch Gusterer et.al. 2007).

Die Erhebung der unabhängigen Variablen muss durch einen entsprechenden Lehrer/innenfragebogen erfolgen. Einzelne unabhängige Variablen (z. B. Teilnahme an Wettbewerben) können auch in der Präambel des Schüler/innenfragebogens von den Schülerinnen und Schülern erhoben werden. Bei der Datenerhebung ist weiters zu beachten, dass die mittels Schüler/innenfragebögen ermittelten Daten mit jenen von den Lehrerinnen und Lehrern anzugebenden Daten für die weiteren Berechnungsverfahren zusammengeführt werden müssen. Dazu ist es notwendig, ein Erhebungsverfahren zu entwickeln, um die Lehrer/innendaten den entsprechenden Schüler/innendaten zuordnen zu können. Die Verfasserin schlägt hierzu vor, die jeweilige Anzahl von Schüler/innenfragebögen in einem Umschlagbogen zu sammeln, welcher inwendig die von den Lehrerinnen und Lehrern erhobenen Daten aufweist. Diese Vorgangsweise ermöglicht eine Zuordnung der Daten bei weitgehender Wahrung der Anonymität der Evaluations Teilnehmer/innen.

Zur Darstellung von Effekten (Wirkungen) unabhängiger Variablen auf die abhängigen eignet sich das Auswertungsverfahren der multivariaten Varianzanalyse. Laut Bühl/Zöfel (1998, S. 380) liegt sie dann vor, wenn der Einfluss von Faktoren und gegebenenfalls Kovariaten (= unabhängigen Variablen) auf mehrere abhängige Variablen in derselben Varianzanalyse gleichzeitig untersucht wird. Nach Bortz/Döring (2002, S. 545 f.) werden bei der multivariaten Varianzanalyse wechselseitige Beziehungen der abhängigen Variablen untereinander berücksichtigt bzw. aufgedeckt. Dies ist bei einer komplexen, abhängigen und nur durch mehrere operationalisierte Indikatoren zu erfassenden Variable äußerst hilfreich.

Allerdings lassen sich bei Hypothesen ohne vorab definierter Effektgrößen auf Grund von Signifikanzen zwar Effekte, nicht aber die Dimensionen dieser Wirkungen nachweisen. Wenn es daher, wie es hier der Fall ist, um eine erwünschte Qualitätsentwicklung in der Begabungsförderung geht, erweitert sich durch die Formulierung von Effektgrößen (d. h.

---

<sup>3</sup> Während das Klassenlehrer-Prinzip es in der Volksschule bei relativ geringem Aufwand ermöglicht, den Unterricht in seiner Gesamtheit zu evaluieren, ist eine summative Evaluation im Sekundarbereich nur in einzelnen Unterrichtsfächern bzw. nach Abschluss einzelner Unterrichtsprojekte vertretbar.

<sup>4</sup> Innere Differenzierung/Individualisierung wurde den abhängigen Prozessvariablen des Schüler/innenfragebogens zugeordnet.

in welchem Ausmaß eine Veränderung stattgefunden haben muss, um von Qualitätsverbesserung zu sprechen) der Interpretationsrahmen der Ergebnisse zu Gunsten einer tatsächlichen Brauchbarkeit (vgl. Urban 2004, S. 38f.).

Eine Aussage zur Wirkung der unabhängigen Variablen auf die abhängigen lässt sich auch durch den Eta-Koeffizienten treffen. Es handelt sich dabei um den Korrelationskoeffizienten, der die linearen und nicht linearen Zusammenhänge zwischen den abhängigen und unabhängigen Variablen erfasst (Bortz 1999, S. 269f.). Der Vorteil liegt in dem möglichen Nachweis einer gegebenenfalls tatsächlich linearen Wirkung, die Rückschlüsse für weitere Qualitätsentwicklungsprozesse zulässt.

Eine Standardisierung (z-Transformation) der Variablenwerte ermöglicht die Vergleichbarkeit unterschiedlicher Stichproben (Bortz 1999, S. 45). Dies ist dann von Nutzen, möchte man einzelne Evaluationsergebnisse im zeitlichen Längsschnitt vergleichen. Auch die Vergleichbarkeit der Ergebnisse unterschiedlicher Schulstandorte wird dadurch möglich. Da in der Standardnormalverteilung der Mittelwert bei 0 liegt und die Variablenwerte in einem einheitlichen Bereich zwischen  $-0,3$  und  $+0,3$ , ist es darüber hinaus relativ einfach, Niveaustufen hoher, mittlerer und niedriger Skalenwerte festzulegen.

### **Ausblick**

Im Sinne der wissenschaftlichen Absicherung des vorgestellten Qualitätsanalyse-Instrumentariums fehlen Untersuchungen zu dem Einfluss der Persönlichkeitsvariablen auf die Prozess- und Produktbeurteilung bei Schülerinnen und Schülern im Pflichtschulalter. Hier müssen die Ergebnisse der Untersuchungen von Urban (2004) an Studierenden herangezogen werden. Es ist aber zu bedenken, dass bei Schülerinnen und Schülern v. a. im Volksschulalter von einer stärkeren emotionalen Bindung an die Lehrpersonen und somit höheren Sympathiewerten ausgegangen werden muss. Inwieweit dies das Beurteilungsverhalten beeinflusst, ist dzt. noch offen.

Darüber hinaus erfüllt der vorgestellte Fragebogen zwar die Forderungen der klassischen Testtheorie nach Objektivität, Reliabilität und Validität, jedoch wurde nicht überprüft, ob er auch den Qualitätskriterien der probabilistischen Testtheorie<sup>5</sup> entspricht. In diesem Modell ergibt sich die Wahrscheinlichkeit einer richtigen Itemlösung als Funktion zweier Modellparameter, nämlich der Fähigkeit der antwortenden Person und der Schwierigkeit des betreffenden Items. Die Zuverlässigkeit und Gültigkeit eines Items ist umso höher, je höher die Vorhersagbarkeit der einzelnen Itemantworten aufgrund der zu messenden Personenvariable ist (Rost 1996, S. 363).

Abschließend wird darauf hingewiesen, dass – bedingt durch die Komplexität eines Ausbildungssystems – es in der Konzeption eines Evaluationsinstruments nicht gelingen kann, alle Aspekte zu entdecken, die für einen Evaluationsprozess wesentlich wären. Erst im wiederholten Einsatz können Defizite, falsche Funktionalitäten und vieles mehr sichtbar werden und Verbesserungen oder Ergänzungen vorgenommen werden (Urban 2004, S. 270).

### **Literatur**

- Altrichter, H., Messner, E. & Posch, P. (2006). Schulen evaluieren sich selbst. Ein Leitfaden. Seelze: Kallmeyer.
- Berendt, B. (1994). „Gute Lehre“ und ihre Planung im Überblick. Aspekte und Fragestellungen zur Bestimmung von Kriterien „guter Lehre“. In Handbuch Hochschullehrer: Informationen und Handreichungen aus der Praxis für die Hochschullehre (S. 1-20). Bonn.

---

<sup>5</sup> Im Rahmen der vorliegenden Pilotstudie hatte die Verfasserin keinen Zugang zu dem entsprechenden Computerprogramm, mit dem eine auf der probabilistischen Testtheorie aufbauenden Rasch-Skalierung (Hanisch 1990) gerechnet werden kann.

- Bortz, J. (1999). Statistik: Für Sozialwissenschaftler. Berlin: Springer.
- Bortz, J. & Döring, N. (2002). Forschungsmethoden und Evaluation. Berlin: Springer.
- Bühl, A. & Zöfel, P. (1998). SPSS Version 8. Einführung in die moderne Datenanalyse unter Windows. München: Pearson Studium.
- Gruehn, S. (2000). Unterricht und schulisches Lernen: Schüler als Quellen der Unterrichtsbeschreibung. (= 12. Band der Reihe pädagogische Psychologie und Entwicklungspsychologie, hrsg. v. D.Rost). Münster: Waxmann.
- Gusterer, W. et.al. (2007). Begabungssiegel. Info-Paket. Wien: Stadtschulrat für Wien.
- Hanisch, G. (2004). Messung von Schulangst. Erziehung und Unterricht. Österreichische pädagogische Zeitschrift, 9-10, S. 897-902.
- Hanisch, G. (1990). Rasch-Skalierung. In E. F. Kleiter (Hrsg.), Lehrbuch der Statistik in KMSS. Band 1/2 (S. 350-388). Weinheim: Deutscher Studienverlag.
- Helmke, A. (2009). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Seelze: Kallmeyer.
- Helmke, A. & Schrader, F. W. (2001). Determinanten von Schulleistung. In D. Rost (Hrsg.), Handwörterbuch Pädagogische Psychologie (S. 81-90). Weinheim: Beltz.
- Österreichisches Zentrum für Begabungsförderung und Begabungsforschung (2006, Hrsg.). Qualitätskriterien für Schulen mit Begabungs- und Begabtenförderung. Salzburg: özbf.
- Rost, J. (1996). Lehrbuch Testtheorie – Testkonstruktion. Bern: Huber.
- Schweer, M. K. W. (1996). Innovation durch Evaluation? Eine kritische Analyse hochschulpolitischer Maßnahmen zur Verbesserung der Qualität der Lehre. Bildung und Erziehung, 49, 158-168.
- Schweer, M. K. W. (2001). Evaluation der Lehre. In D. Rost (Hrsg.), Handwörterbuch Pädagogische Psychologie (S. 159-163). Weinheim: Beltz.
- Urban, W. (2004). Evaluation des tertiären Bildungssystems. Konzeption, Modellbildung und Durchführung. Innsbruck: Studienverlag.
- Weblar, W. D. (1991). Kriterien für die gute akademische Lehre. Das Hochschulwesen, 6, 243-249.
- Wottawa, H. & Thierau, H. (1990). Lehrbuch Evaluation. Bern: Huber.

Anm.: Die vorliegende Pilotstudie wurde nach Absprache mit Herrn BSI Walter Gusterer, MSc entwickelt. Die Verfasserin möchte sich an dieser Stelle herzlich für die tatkräftige Unterstützung bedanken.

Der im Beitrag vorgestellte Schüler/innenfragebogen zur Bewertung von Unterrichtsqualität wird dem Kompetenzzentrum für Begabungsförderung des SSR für Wien zur weiteren Verwendung zur Verfügung gestellt.

DR. GUNDULA WAGNER  
 Teamlehrerin für Begabtenförderung  
 Volksschule Pfeilgasse, Wien 8  
[gundula.wagner@gmx.at](mailto:gundula.wagner@gmx.at)