

Christian Liesen und Marco G. P. Hessels

Zur Brauchbarkeit der P-Scales für standardbasierte Kompetenzmessung in der Schweiz

Zusammenfassung

Standardisierung und sonderpädagogische Förderung – passt das zusammen? Instrumente wie die englischen P-Scales (vorgestellt in SZH Nr. 1/2008) bejahen das. Sie versprechen, den Lernstand entwicklungsbezogen zu erfassen und zu objektivieren. Eine deutsche Übersetzung wurde Lehrpersonen vorgelegt und sie wurden gebeten, die Brauchbarkeit einzuschätzen.

Résumé

Standardisation et renforcement de la pédagogie spécialisée – ces concepts sont-ils conciliables? Des instruments, comme le «P-Scales» élaboré en Angleterre et présenté dans la revue en langue allemande N° 1/2008 du CSPS), l'affirment. Ces instruments promettent d'évaluer de manière objective les stades d'apprentissage par rapport au développement. Une traduction allemande a été proposée au corps enseignant qui est prié d'en analyser sa faisabilité.

Seit Mitte des Jahres hat die Schweiz nationale Bildungsziele – HarmoS machte es möglich. Jetzt ist nachzulesen, welche Grundkompetenzen in Sprachen, Mathematik und Naturwissenschaften den «Kern der schulischen Bildung» ausmachen. Das Niveau wurde empirisch bestimmt: Rund 12 000 Schülerinnen und Schüler erhielten Testaufgaben. Wenn 80–90 Prozent (und unter keinen Umständen weniger als 2/3) von ihnen sie lösen konnten, ergab dies das Grundniveau (vgl. EDK, 2011, S. 4). Per Definition verfügen also die meisten Schülerinnen und Schüler bereits über den Kern der schulischen Bildung. Ziel der Übung ist laut EDK nicht ein Kompetenzzuwachs, sondern die Harmonisierung der Bildungsziele: Es wird erwartet, dass sich Lehrmittel, Lehrpläne und Evaluationsinstrumente künftig stärker an den Grundkompetenzen ausrichten (vgl. www.edk.ch/dyn/12930.php).

Aus sonderpädagogischer Sicht stellt sich die Frage: Nicht alle Schülerinnen und Schüler werden diesen Minimalstandard er-

reichen. Was ist mit ihnen? 2008 haben Luder, Prammer und Prammer-Semmler in dieser Zeitschrift von den englischen P-Scales berichtet. Es handelt sich dabei um eine Sammlung von Skalen zur standardbasierten Kompetenzmessung. In England wird sie immer dann angewandt, wenn eine Schülerin oder ein Schüler nicht an den dort obligatorischen, landesweiten Leistungstests teilnehmen kann, weil sie oder er das definierte Mindestniveau dieser Leistungstests nicht erreicht. Mit den P-Scales soll eine zuverlässige Einschätzung des Lernstands auch in diesen Fällen möglich sein. Inzwischen sind die P-Scales offizieller Teil des englischen nationalen Curriculums (Ndaji & Tymms, 2009a, S. 18).

Aufbauend auf den Vorarbeiten von Luder et al. formierte sich 2009 eine kleine Arbeitsgruppe, um die Frage der standardbasierten Kompetenzmessung aus sonderpädagogischer Sicht nicht aus den Augen zu verlieren. Als Projektkronym bürgerte sich SECEL ein, Skalen zur Entwicklungs-

und Curriculumbasierten Einschätzung des Lernstands. Getragen wird die Gruppe von Reto Luder (PHZH), Marco Hessels (Universität Genf), Christian Liesen (HfH) und dem SZH (massgeblich René Stalder). Das Ziel ist herauszufinden, ob und wie sich individuelle Entwicklung und individueller Lernstand gemeinsam betrachten lassen, wenn standardbasierte Skalen zugrundegelegt werden.

Die bisherige Arbeit der Gruppe konzentriert sich auf die P-Scales. Es wurden Übersetzungen in deutscher und französischer Sprache angefertigt. Die deutsche Übersetzung wurde in Zürich jetzt erstmals Schweizer Lehrpersonen vorgelegt – davon handelt der folgende Bericht. Die französische Übersetzung wird in Genf in einer noch laufenden Studie eingesetzt. Nicht unerwähnt soll bleiben, dass die P-Scales in England unterdessen Gesellschaft bekommen haben: Es gibt mindestens fünf Implementierungen, die alternativ eingesetzt werden (sie hören auf die Namen PACE/PACE2, PIVATS, B-squared, Green Box und CASPA, vgl. Ndaji & Tymms, 2009b, S. 183–189). Um «Implementierungen» handelt es sich, weil alle fünf Instrumente auf den P-Scales beruhen. Die P-Scales sind ja Teil des nationalen Curriculums, ihr Grundgerüst bleibt deswegen in allen fünf Fällen erhalten. Sie dienen als Vorlage und Referenzrahmen. Doch wo die P-Scales acht hierarchische Leistungsstufen unterscheiden (P1 bis P8) wovon fünf fächerspezifische Lernleistungen beinhalten (P1 bis P3 behandeln basale Lernfähigkeiten, P4 bis P8 sind differenziert nach Sprache, Mathematik und Naturwissenschaften), weichen die genannten Implementierungen davon teilweise ab. Sie organisieren und gewichten die Leistungsstufen anders und mischen Erweiterungen bei.

Diese Entwicklung ist sehr interessant. Über kurz oder lang wird sie auch im Projekt SECEL ihren Niederschlag finden. Noch aber gilt den P-Scales exklusiv die Aufmerksamkeit. Ausführlich dargestellt werden sie in Luder et al. 2008, auf das hier statt näherer Ausführungen nochmals hingewiesen sei.

Eine Interviewerhebung zum Sinn und Nutzen von Standardisierung

Die P-Scales dienen denn auch als Gegenstand für eine Untersuchung, die danach fragte, wie Heilpädagoginnen und Heilpädagogen zu standardisierten Kompetenzmodellen stehen. Es ist oft zu hören, dass durch den Einsatz standardisierter Instrumente eine bessere Schulqualität erreicht werden kann. Aus heilpädagogischer Sicht stellen sich jedoch Sinn und Nutzen, Vorteile und Nachteile von Standardisierung nicht ohne weiteres so glatt dar.

Eine Studentin der HfH war glücklicherweise abenteuerlustig genug, ihre Masterthese in diesem Bereich zu verfassen (Böni, 2011)¹. Dadurch liegen nun erstmals Antworten zu folgenden beiden Fragen vor:

1. Welche Themen beschreiben praktisch tätige Heilpädagoginnen und Heilpädagogen hinsichtlich der Brauchbarkeit von Standardisierungsinstrumenten?
2. Wie brauchbar schätzen sie die P-Scales ein?

Zu diesen Fragen wurde ein Interviewleitfaden ausgearbeitet. Er umfasste vier Themenbereiche: (1) allgemeine Einstellung zu Standardisierung, (2) Nutzen der P-Scales, (3) Klarheit und Anwendungsfreundlichkeit

¹ Wir danken Monika Böni für ihren Beitrag zum Projekt SECEL und für ihr Einverständnis, ausgewählte Punkte auch in diesem Format zu publizieren.

der P-Scales, (4) Kenntnis und Einschätzung anderer standardbasierter Instrumente. Mit insgesamt 15 Heilpädagoginnen und Heilpädagogen, beschäftigt im Kanton Zürich, wurden auf dieser Grundlage teilstrukturierte Leitfadeninterviews geführt. Sie wurden anschliessend vollständig transkribiert und mit MAXQDA 10 inhaltsanalytisch ausgewertet. Um Differenzierungsmöglichkeiten zu gewinnen, war die Gruppe der Interviewten aus je fünf IF-, IS- und HPS-Lehrpersonen zusammengesetzt, so sind sowohl hoch- als auch niederschwellige Angebote in der Stichprobe vertreten². Die Ergebnisse der Auswertung wurden in einem Fokusgruppeninterview validiert.

Ergebnisse

Allgemeine Einstellung zu Standardisierung
Eines der wichtigsten Ergebnisse der Studie ist, dass Lehrpersonen zwei verschiedene Dinge vor Augen haben, wenn sie von Standardisierung sprechen.

1. Standardisierte Instrumente (Tests) werden als eher positiv beurteilt. Das professionelle Anwenden von standardisierten Tests kann demnach sinnvoll sein, weil es einen Vergleich mit einer Norm / Richtlinie ermöglicht, sich gezielt Hinweise für die Förderplanung ergeben können und weil die interdisziplinäre Zusammenarbeit vereinfacht wird, nicht zuletzt durch eine gemeinsame Sprache. Dabei gibt es über alle drei Befragungsgruppen hinweg ein ausgeprägtes Bewusstsein dafür, dass mit standardisierten Instrumenten

nur das beobachtet und bewertet wird, was das Instrument vorsieht. Mit anderen Worten, bestimmte Kompetenzbereiche werden hervorgehoben, während andere vernachlässigt werden.

2. Bildungsstandards als solche werden eher negativ beurteilt. Vor allem die IF-Lehrpersonen äussern sich dazu, weil sich für sie die Frage am dringendsten stellt: Sie unterstützen Kinder im Regelunterricht, ohne dass die Sonderschulung involviert ist. Entsprechend stärker scheint sich die Orientierung an den Zielen und Vorgaben der Regelschule auszuwirken, nach dem Motto: Wir müssen im Unterricht das Gleiche für alle machen. IS- und HPS-Lehrpersonen sind demgegenüber der Auffassung, dass ihre Schülerinnen und Schüler – die in beiden Fällen der Sonderschulung unterstellt sind – die Standards ohnehin nicht erreichen können. Sie ignorieren sie weitgehend.

Verschiedene Interviewteilnehmer sagen, dass eine standardbasierte Kompetenzmessung gefährlich sein kann. Sie erwähnen die Gefahr, dass nur das gemessen wird, was leicht zu messen ist, und sie befürchten verstärktes «Teaching to the test», also die einseitige Ausrichtung und Gestaltung von Unterricht an Testvorgaben. Sie geben ausserdem zu bedenken, dass sich ein offener oder verdeckter Missbrauch der Ergebnisse nicht ausschliessen lässt – dabei denken sie an Schülerbewertungen oder Ranglisten. Schliesslich wird noch hervorgehoben, dass das Ergebnis der Messung oft von der Lehrperson abhängig sei.

² IF: Integrative Förderung, d.h. eine den Unterricht in der Regelklasse ergänzende Förderung ohne Sonderschulstatus. – IS: Integrierte Sonderschulung, d.h. Schulung im Klassenverband der Regelschule, aber mit individuellen Zielen und administrativ der Sonderschulung zugeteilt. – HPS: Lehrperson an einer heilpädagogischen Schule.

Nutzen der P-Scales

Der Inhalt der P-Scales wird insgesamt als differenziert und brauchbar eingeschätzt. Dafür sind offenbar vor allem die zahlreichen Beispiele verantwortlich, mit denen die Leistungsstufen und die einzelnen Items darin illustriert werden. Auch der logische und stringente Aufbau werden gelobt. Fast alle befragten Lehrpersonen können sich vorstellen, durch die Anwendung der Skalen eine bessere Basis für die interdisziplinäre Zusammenarbeit zu erhalten: Dafür sei ein gemeinsames Vokabular nützlich, wie die P-Scales es bieten. So könne die Verständigung über den Entwicklungs- und Leistungsstand eines Kindes über Berufsgrenzen hinweg erleichtert werden.

Kritisch wird eingewandt, dass die Einschätzung des Entwicklungsstandes eines Kindes immer eine grosse Aufgabe und Herausforderung sei. Der Einsatz standardbasierter Instrumente mag dazu einen Beitrag leisten. Die Instrumente blieben letztlich aber an der Oberfläche – auch ein gut durchdachtes wie die P-Scales.

Insgesamt schält sich aus den Antworten ein Katalog an Wunschkriterien heraus, die ein brauchbares, standardbasiertes Instrument aus Sicht der drei befragten Gruppen vereinen sollte.

Es muss:

- verständlich und übersichtlich sein (auch im Hinblick auf verschiedene Berufsgruppen)
- praktikabel sein, d. h. eher weniger Text zugunsten von bspw. Checklisten enthalten
- ansprechend gestaltet sein
- nicht viel Zeit für die Anwendung verlangen (30 Minuten)
- beschreibend, nicht beurteilend sein.

Klarheit und Anwendungsfreundlichkeit der P-Scales

Kritikpunkte an den P-Scales gab es natürlich auch. Die meisten finden sich in diesem Themenbereich. So war ein immer wieder vorgebrachter Punkt das Layout der deutschsprachigen Übersetzung. Es entsprach nicht näherungsweise der Aufmachung des englischen Originals, sondern bestand aus langen Texteinheiten³. Dies erfüllte seine Funktion für den Zweck der Studie, ist von einer praxisfreundlichen Gestaltung aber weit entfernt. Vor allem die IF-Lehrpersonen hatten Schwierigkeiten damit.

Ein weiterer Punkt betraf die Differenziertheit der Leistungsstufen. Die Stufen P1 bis P3, die basale Lernfähigkeiten behandeln, wurden als zu wenig kleinschrittig eingeschätzt. Die Stufen P7 und P8 entsprechen nicht dem Übergang vom Kindergarten in die Grundschule in der Schweiz, obwohl sie dies abbilden können sollten. Eine feinere Unterteilung für Schwerbehinderte sei grundsätzlich auf allen Stufen P1 bis P8 notwendig. Auffallend ist, dass alle diese Punkte im Zusammenhang mit der Anwendungsfreundlichkeit der P-Scales genannt wurden und nicht im Zusammenhang mit ihrem Nutzen.

Kenntnis und Einschätzung anderer standardbasierter Instrumente

Zu diesem Punkt werden eine Reihe von Instrumenten genannt, die Leistungs- und Lernstandseinschätzungen in vereinheitlichender, übergreifender Weise erlauben. Dabei handelt es sich nicht um standardbasierte Instrumente im eigentlichen Sinne, son-

³ Es handelte sich um die interne Arbeitsübersetzung der SECEL-Projektgruppe.

dern um klassifikatorische oder prozessuale Instrumente. Genannt werden im IF-Bereich namentlich das LernLOT, das Cockpit und der Instrumentenkoffer von St. Gallen, ferner schuleigene Kompetenzraster und selbst kreierte Beobachtungsbögen. Im IS- und HPS-Bereich werden ICF-basierte Förderplanner, der Webbasierte Förderplanner (WFP) und der Förderdiagnostische Assistent (FDAss) genannt. Flächendeckend eingesetzte Instrumente gibt es nicht. Dies wird in den Interviews immer wieder als Problem für eine bessere Verständigung über Lern- und Leistungsstände angesprochen. Das Standardisierte Abklärungsverfahren (www.sav-pes.ch) findet noch keine Nennung.

Diskussion und Ausblick

Was also ist mit den Schülerinnen und Schülern, die den Minimalstandard nicht erreichen? Viele der Befragten können sich vorstellen, eine standardbasierte Skala praktisch zu benutzen. Die insgesamt deutlich positive Einschätzung der P-Scales lässt darauf schliessen, dass ihr inhaltlich durchdachter Aufbau auch hiesige Lehrpersonen anzusprechen vermag: Ziele unterhalb der minimalen Bildungsziele sind denkbar.

Über die Schwierigkeiten kann das freilich nicht hinwegtäuschen. Drei davon seien abschliessend genannt.

- Standardisierung ist für die Heilpädagogik ein ungewohntes Thema. Das ist in den Äusserungen der Befragten gut spürbar. Immer wieder scheint Unsicherheit bezüglich der Konzepte und Begriffe durch.
- Es tritt ein Widerspruch zwischen Standardisierung und Individualisierung zutage. Individualisierung gehört zum professionellen Selbstverständnis. Alle

Befragten lassen erkennen, dass sie nicht glauben, dass ein standardisiertes Instrument der Lern- und Leistungsentwicklung der Kinder «wirklich» gerecht werden kann. Sie lassen auch erkennen, dass sie nicht glauben, dass es «wirklich» zur Verbesserung der Schulqualität beiträgt – entscheidend sei immer noch der Unterricht. Das Bild, das man durch eine standardbasierte Brille sieht, ist für die Befragten nur ein Ausschnitt aus der Lebens- und Entwicklungsrealität des Kindes. Die Crux dabei ist aber, dass nicht die Individualisierung, sondern die Standardisierung vergleichsweise sicheren Boden anzubieten scheint für professionelle Urteile. Ist alles andere nicht latent dem Vorwurf subjektiver Beliebigkeit ausgesetzt? Beginnt man sich nicht auf einmal zu rechtfertigen für die Individualisierung, die nicht ins Schema passt? Das Unbehagen gegenüber der Deutungsmacht, die standardbasierte Leistungsmessungen mit sich bringen, ist deutlich vernehmbar.

- Ein dritter Punkt betrifft die Vorstellung von Entwicklung, die den P-Scales zugrundeliegt. Entwicklungstheorien sind reichhaltig (klassisch: Flammer, 2009). Standardbasierte Leistungsmessung führt aber fast selbstverständlich zu einer Bevorzugung von Entwicklungsvorstellungen, die universell, linear, hierarchisch und stufenorientiert aufgebaut sind. Doch nichts daran ist zwangsläufig. In der Studie gaben mehrere der Befragten ihrer Überzeugung Ausdruck, dass kindliche Entwicklung nicht linear verlaufe; insbesondere seien Rückschritte möglich. Hinzu kommt, dass Standardisierungen unerwünschte neue Kategorisierungen schaffen können.

Solche tiefer greifenden inhaltlichen Fragen lassen erkennen, dass es in Sachen standard-basierter Erfassung von Entwicklung noch viel zu tun gibt. Das gilt auch hinsichtlich des Fernziels des SECEL-Projekts, individuelle Entwicklung und individuellen Lernstand gemeinsam zu betrachten, also gerade die Schnittstelle zwischen Entwicklung und Curriculum.

Eine deutschsprachige Version der P-Scales braucht zweifelsohne noch Feinschliff. Die skizzierten Probleme für den praktischen Einsatz werden sich damit aber nicht aus der Welt schaffen lassen.

Im Bewusstsein dieser Umstände waren die Resultate der Böni-Studie jedenfalls ertragreich genug, um eine erste praktische Erprobung beginnen zu können. Eine zweite Masterthese dazu ist an der HfH bereits in Arbeit. Zudem ist schon absehbar, dass die Genfer Studie, die derzeit mit französischsprachigen Lehrpersonen läuft, wichtige weitere Einsichten liefern und einige Lücken schliessen wird. Die Projektgruppe SECEL wird in loser Folge über den Fortgang der Arbeiten berichten.

Prof. Dr. Christian Liesen
Hochschule für Heilpädagogik
Zürich
Schaffhauserstrasse 239
8057 Zürich
christian.liesen@hfh.ch



Prof. Dr. Marco G. P. Hessels
Universität Genf
Fakultät Psychologie und
Erziehungswissenschaften
40, Bd Pont D'Arve
1205 Genf
Marco.Hessels@pse.unige.ch



Literatur

- Böni, M. (2011). *Interviewerhebung über die Brauchbarkeit von Standardisierungsinstrumenten im heilpädagogischen Bereich – dargestellt am Beispiel der P-Scales*. Masterthese im HfH-Studien-gang Pädagogik bei Schulschwierigkeiten. Inter-net: www.bscw-hfh.ch/pub/bscw.cgi/d5751483/BöniMAT.pdf [Stand 11.8.2011]
- Flammer, A. (2009). *Entwicklungstheorien. Psycholo-gische Theorien der menschlichen Entwicklung* (4., vollst. überarb. Aufl.). Bern: Huber.
- Luder, R., Prammer, W. & Prammer-Semmler, E. (2008). Standardbasierte Kompetenzmessung un-terhalb von Mindestanforderungen: ein Beispiel aus England. Die «P-Scales» und ihre Anwendung im deutschsprachigen Raum. *Schweizerische Zeit-schrift für Heilpädagogik*, 14 (1), 21–31.
- Ndaji, F. & Tymms, P. (2009a). The National Curricu-lum and the Development of the P scales. In Ndaji, F. & Tymms, P. (Hrsg.), *The P Scales. Assessing the Progress of Children with Special Educational Needs* (S. 1–23). West Sussex: Wiley.
- Ndaji, F. & Tymms, P. (2009b). Ways forward. In Nda-ji, F. & Tymms, P. (Hrsg.), *The P Scales. Assessing the Progress of Children with Special Educational Needs* (S. 181–194). West Sussex: Wiley.
- Schweizerische Konferenz der kantonalen Erzie-hungsdirektoren (EDK) (2006). *Faktenblatt: Nati-onale Bildungsziele für die obligatorische Schule: in vier Fächern zu erreichende Grundkompetenzen*. Internet: www.edudoc.ch/static/web/arbeiten/harmos/grundkomp_faktenblatt_d.pdf [Stand 4.8.2011]