

Peter BENDER, Paderborn (in: Beiträge zum MU 2004, Franzbecker)

Die etwas andere Sicht auf PISA sowie TIMSS und IGLU

(gekürzte und weiter entwickelte Version des Beitrags "Die etwas andere Sicht auf die internationalen Vergleichs-Untersuchungen TIMSS, PISA und IGLU" in den Paderborner Universitätsreden 89 (2003), 35-59, im Internet zu finden unter "math-www.uni-paderborn.de/~bender/")

0. Einleitung

Mein Favorit unter all den Konsequenzen, die aus PISA gezogen wurden (diese: vom ehemaligen Ministerpräsidenten eines großen Bundeslandes und inzwischen Mitglied der Bundesregierung): PISA habe gezeigt, dass die Schulzeit bis zum Abitur von 13 auf zwölf Jahre gesenkt werden müsse.

Wegen der beschränkten Zeit konzentrierte ich mich im Vortrag auf mathematik-didaktische Fragen i.e.S. und fasste einen Großteil der in den Paderborner Universitätsreden durchgeführten Analyse in Thesen zusammen:

These 1: IGLU (PISA, TIMSS usw. entsprechend) misst nicht *die* Mathematik- (Lese-, naturwissenschaftlichen usw.) "Leistungen am Ende der 4. Jahrgangsstufe im internationalen Vergleich", sondern die Leistungen bei *diesem* Test unter *diesen* Bedingungen, die, wie bei jedem Test, zu einem großen Teil darin bestehen, das Gemeinte zu entschlüsseln.

These 2: Es steht dahin, ob (a) dem im Angelsächsischen entwickelten Konstrukt der Mathematical Literacy (ML) in der dort festgelegten Ausprägung wirklich *die* fundamentale Rolle im MU zukommen soll, (b) die Test-Aufgaben für dieses Konstrukt valide sind, (c) es seriös ist, die Länder-Curriculum-Validität zu ignorieren und zugleich Länder-Punktzahl-Vergleiche zu veröffentlichen, (d) bei noch so guten Übersetzungen nicht doch diejenigen Kinder es leichter haben, in deren Mutter-Sprache die Aufgaben ursprünglich formuliert worden sind, nämlich in Englisch.

These 3: Es steht dahin, ob die statistischen und organisatorischen Vorgaben wirklich weltweit hinreichend skrupulös eingehalten wurden.

These 4: (a) Für die deutsche Öffentlichkeit waren die schockierenden Ergebnisse heilsam, da die Relativierungen 1.–3. nichts an den Grund-Tendenzen ändern. (b) Für Jemanden mit Einblick in die Schul-Realität waren sie keine Überraschung: Wer wissen wollte, konnte auch schon vorher wissen. (c) Hoffentlich wird der ganze Aufwand durch den Erfolg von in die Wege geleiteten und noch zu leitenden Maßnahmen gerechtfertigt. Die bisherige Diskussion in Deutschland gibt eher zu Skepsis Anlass.

These 5: (a) Viel mehr als über das Bildungs-System sagen die Statistiken jedoch etwas über gewisse Mangel-Erscheinungen unserer Gesellschaft

aus, besonders die schlechte Integration der Kinder mit Migrations-Hintergrund (mindestens ein Elternteil nicht in Deutschland geboren), vor allem wenn *beide* Elternteile im Ausland geboren sind. Diese Mangel-Erscheinungen können nicht allein innerhalb des Bildungs-Systems repariert werden, wenn überhaupt. (b) Die gesellschaftlich-kulturellen Bedingungen werden unterschätzt, z.B. der Einfluss der (aus unserer Sicht vielleicht abzulehnenden, aber eben TIMSS-, PISA- und IGLU-erfolgreichen) Leistungs- und Disziplin-Orientierung in vielen ost-asiatischen Ländern oder deren Verlust in den früheren sozialistischen Ländern in Verbindung mit dem Verlust der autoritären Strukturen dort. Dagegen sind z.B. die aus der TIMSS-Video-Studie (deren Repräsentativität außerdem natürlich in Frage steht) extrahierten Unterschiede in den Unterrichts-Stilen in Japan, USA und Deutschland so gut wie belanglos. (c) Die Vergleichs-Studien geben weder etwas für noch gegen eine Bevorzugung der Einheits-Schule her.

1. Das Idealbild eines Tests mathematik-bezogener Kompetenzen und die Realität

Welche mathematik-bezogenen Fähigkeiten, Fertigkeiten, Wissens-Bestände, Einstellungen usw. hält man für wichtig, so dass man mit dem Grad ihres Vorhandenseins mathematische Leistungs-Fähigkeit eines Individuums oder einer ganzen Population bestimmt? Wie misst man diese Tugenden? Üblicherweise lässt man übliche Aufgaben lösen, und zwar bei einer großen Population überwiegend solche, bei denen die Antwort entweder richtig oder falsch ist. Dabei unterstellt man Validität, d.h. dass tatsächlich die interessierenden Tugenden relevant sind. Allerdings gibt es dazu keine robusten mathematik-didaktischen Forschungs-Ergebnisse.

Es ist klar, dass in einem solchen Test viele wichtige Tugenden nicht berücksichtigt werden: Die Fähigkeit, komplexe Probleme anzugehen, überhaupt Mathematisierbarkeit zu prüfen, ein Problem längerfristig zu bearbeiten, Ansätze zu verwerfen oder weiter zu verfolgen, das Problem eine Zeit lang liegen zu lassen, Anderen es verständlich darzustellen, von Gesprächen mit Anderen zu profitieren, Medien inklusive Internet zu nutzen, usw.

2. Mehr oder weniger gute Aufgaben-Formulierungen

3. Was testen die Aufgaben eigentlich wirklich ab?

Das von den Aufgaben-Autorinnen & -Autoren Gemeinte herauszufinden, wird besonders bei außer-mathematischen Kontexten, misslungenen Formulierungen bzw. schwachen Übersetzungen schwierig. Diejenigen sind im Vorteil, die an solche (standardisierten, von Fremden gestellten) Tests gewöhnt sind und in deren Sprache und Kultur die Aufgaben ursprünglich angesiedelt sind.

4. Das "ML"-Konzept auf Kosten der curricularen Validität

Während bei TIMSS die Validität bezüglich der Länder-Curricula noch ein erklärtes Ziel war, findet bei PISA ein "Verzicht auf transnationale curriculare Validität" statt, stattdessen wird "ein didaktisches und bildungstheoretisches Konzept mit sich [geführt], das normativ ist". Als hierfür "in mancher Hinsicht vorbildlich" werden die NCTM-Standards herangezogen. Es ist klar, dass Länder, die ihr geschriebenes und reales Curriculum daran ausgerichtet haben, ob durch explizite Übernahme oder durch eigene Entwicklung, von PISA "bevorzugt" werden. — Wie der Mehrheit in der deutsch-sprachigen mathematik-didaktischen Kommunität sagt mir das ML-Konzept von PISA für den MU durchaus zu, und ich teile die Einschätzung schon der TIMSS-Leute, dass der deutsche MU allzu sehr auf Faktenwissens-Erwerb und die Beherrschung von Verfahren zielt.

Definition von ML: "Die Rolle zu erkennen und zu verstehen, die die Mathematik in der Welt spielt, fundierte mathematische Urteile abzugeben und sich auf eine Weise mit der Mathematik zu befassen, die den Anforderungen des gegenwärtigen und künftigen Lebens einer Person als konstruktivem, engagiertem und reflektierendem Bürger entspricht."

Diese "Definition" passt durchaus zur Tradition der deutschen bildungstheoretischen Didaktik, wie sie z.B. vom alten Wolfgang Klafki (1958) oder für den MU von Heinrich Winter (1975) verkörpert wird.

Der dominierende Bestandteil von ML im Sinne von PISA ist die Kompetenz zum Modellieren (offenbar i.W. mit Problemlösen und sogar erklärtermaßen mit Aufgaben-Lösen zu identifizieren; I2, 118f) in den Kompetenz-Klassen: (i) Verfahren ausführen; (ii) Zusammenhänge herstellen; (iii) Verallgemeinern. Besonders wichtig unter den NCTM-Standards: "Vorbereitung auf offene Aufgabenstellungen, da realistische Probleme und Aufgaben in der Regel nicht gut definiert sind".

5. Wie weit kann ML überhaupt getestet werden?

Tests wie PISA, TIMSS oder IGLU, zumal mit überwiegend Multiple-Choice-Aufgaben, können der o.a. Definition natürlich nicht gerecht werden (so schon Kießwetter 2002). Kein einziger Aspekt kann sich in solchen Aufgaben genuin wieder finden. Natürlich ist keine wirklich offen; es ist lediglich immer wieder der Versuch erkennbar, ein direktes Anwenden von Fakten-Wissen und Fertigkeiten durch Einkleidung des mathematischen Gehalts in allerlei inner- und außer-mathematische Kontexte zu verhindern und so immerhin Modellieren zu erzwingen. — Vielleicht sind aber solides Wissen und Können doch viel bedeutsamer für die ML, als der NCTM und die PISA-Leute annehmen.

6. Das fragwürdige Konstrukt der Kompetenz-Stufen

Nach den PISA-Tests wurde die Punkte-Skala in fünf Abschnitte je der Länge 91,75 eingeteilt, und diese wurden zu Kompetenz-Stufen erklärt.

V	≥696	Komplexe Modellierung und innermathematisches Argumentieren,
IV	604–695	Umfangreiche Modellierungen auf der Basis anspruchsvoller Begriffe,
III	512–603	Modellieren und begriffliches Verknüpfen auf dem Niveau der Sekundarstufe I,
II	421–511	Elementare Modellierungen,
I	329–420	Rechnen auf Grundschulniveau,
0	≤328.	

Die aufgeprägte (ja eigentlich zugrunde liegende) Leistungs-Stufung beim Modellieren ist inkonsistent (von "nicht" über "elementar", "SI-Niveau", "umfangreich" bis "komplex") und mit nicht-konsistenten und nichtssagenden Zusätzen ("begriffliches Verknüpfen", "anspruchsvolle Begriffe", "innermathematisches Argumentieren") versehen.

Die PISA-Leute kommen zu der Feststellung, dass man erst ab der Kompetenz-Stufe III von einem ausreichenden Standard für ML sprechen sollte (und in Deutschland nur 44% sich mindestens dort befinden), — beruhend auf der o.a. willkürlichen Kompetenzstufen-Setzung.

7. Die Kompetenz-Stufen bei IGLU

8. Resümee

These 6: (a) Aus mathematik-didaktischer Sicht ist das Instrument der Kompetenz-Stufen fragwürdig. Die Einteilung in Stufen ist willkürlich und ersichtlich nicht das Ergebnis einer (mathematikdidaktisch-kognitionspsychologischen) Theorie. (b) Der Versuch, sich von der ausschlaggebenden Bedeutung der konkreten mathematischen Inhalte (und der nationalen Curricula) zu lösen, muss als gescheitert angesehen werden. Die zentralen Konstrukte "ML", "Modellbildung", "Problemlösen", "Anwenden" sind z.T. praktisch nichtssagend, z.T. nicht mit solchen Tests abprüfbar.

Literatur (weitere im Aufsatz unter der eingangs angegebenen Internet-Adresse)

(I2, IGLU) Wilfried Bos, Eva-Maria Lankes, Manfred Prenzel, Knut Schwippert, Renate Valtin, Gerd Walther (Hrsg.) (2004): IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich. Münster u.a.: Waxmann

Kaiser, Gabriele (2000): Internationale Vergleichsuntersuchungen im Mathematikunterricht — eine Auseinandersetzung mit ihren Möglichkeiten und Grenzen. In Journal für Mathematik-Didaktik 21, 171–192