Praktische Anwendungsaufgaben zur integrierten Förderung formaler und materialer Kompetenzen

Erträge aus dem TIMSS-Experimentiertest

Rita Stebler, Kurt Reußer & Erich Ramseier


Der TIMSS-Experimentiertest - auf dem Weg zu einer neuen Prüfungskultur

TIMSS ist das Kürzel für Third International Mathematics and Science Study. In dieser auf der Sekundarstufe I und II durchgeführten Untersuchung wurden weltweit die Mathematik- und Naturwissenschaftsleistungen sowie deren Rahmenbedingungen verglichen. Auf der Sekundarstufe I bestand die TIMSS-Erhebung aus einem schriftlichen Leistungstest, an dem sich 41 Länder beteiligten, und einem Experimentiertest, der in 21 Ländern durchgeführt wurde.


Verstehen und Problemlösen - die Leitideen einer neuen, authentischeren Prüfungskultur


Lange Zeit wurde nicht erkannt, dass solche Prüfungsaufgaben unerwünschte Wirkungen auf die schulischen Lehr-Lern-Kulturen haben können. Um bei den staatlichen Leistungsvergleichen gut abzuscheiden und Lohnneibussen zu vermeiden, richten viele Lehrpersonen ihren Unterricht inhaltlich und didaktisch auf die dort eingesetzten Aufgaben aus (Silver, 1992). Sie vermitteln ihren Schülerinnen und Schülern ein umfassendes Faktkenwissen, das sich im Schulleistungstest zwar auszahl, in Alltag und Beruf aber wenig taugt. Gründliches Verstehen, vernetetes Denken, kreatives Problemlösen und kompetente Selbst-


Experimentieraufgaben lösen - formale und materiale Kompetenzen orchestrieren

Durch den Einbezug anspruchsvoller Anwendungsaufgaben in die Schulleistungstests will man in den USA zum einen Verstehens- und Problemlöseleistungen kontextualisiert und umfassender messen (Baxter & Shavelson, 1994), zum anderen die Unterrichtsqualität und den Lern- und Denkprozesse der Schülerinnen und Schüler stärker in Richtung 'higher order thinking' lenken (Resnick & Resnick, 1992; Shepard, 1992). Im Hinblick auf diese Ziele werden u.a. authentische, lebensweltliche, bedeutsame, interessante, herausfordernde und anspruchsvolle Experimentieraufgaben empfohlen (Shavelson & Baxter, 1992). Man hat dabei komplexe, in fachlichen Kontexten verankerte Handlungsprobleme vor Augen, die relevante Bildungsinhalte verkörpern, mehrere Lösungen zulassen und Lern- und Problemlöseaktivitäten auslösen, die sich mit den in Abbildung 1 dargestellten Attributen charakterisieren lassen.

verstehensbezogen

anwendungsorientiert

authentisches
denkprozessbezogen

phänomenbezogen

handlungsbegründet

induktiv

situativ

Je nach Experimentieraufgabe müssen die Schülerinnen und Schüler ein greifbares Produkt herstellen oder Objekte bezüglich bestimmter Merkmale vergleichen, die Bestandteile eines Ganzen identifizieren, Objekte klassifizieren, einen Vorgang über längere Zeit systematisch beobachten und protokollieren sowie durch logisches Denken und vernünftiges Urteilen begründete Schlussfolgerungen ziehen (Shavelson, et al. 1997). Es wird erwartet, dass sie sich wie Forscher (Shavelson, 1996) und Fachexpert*innen (Baxter, Elder, & Glaser, 1996) verhalten, die Hypothesen generieren, planen, prüfen, überwachen, Daten protokollieren, beurteilen und die Ergebnisse anhand ihres Fachwissens erklären.

Im Kontext von TIMSS werden diese und weitere aufgabenübergreifende Tätigkeiten als Leistungserwartungen (performance expectations) bezeichnet. Die Funktion dieses messtheoretischen Konstruktiv ist es, to describe, in a non-hierarchical scheme, the many kinds of performances or behaviours that a given test item or block of content might be expected to elicit from students’ (Robitaille, et al., 1993, p. 44). Diese Umschreibung verweist auf die Tatsache, dass man kognitive Kompetenzen nicht direkt, sondern nur über Verhaltensauswirkungen erfassen kann. Wenn man also denken und Problemlösen messen will, muss man als erstes jenes Verhalten beschreiben, von dem man annimmt, dass es Rückschlüsse auf die zu erfindenden Kompetenzen zulässt. Als nächstes muss man Testaufgaben entwickeln, die das gewünschte Verhalten auslösen. Schließlich müssen Experten beurteilen, wie sehr die Aufgaben und die durch sie provozierten Verhaltensauswirkungen inhaltlich mit dem vorher beschriebenen Verhalten übereinstimmen (Ingkamp, 1985; Gage & Berliner, 1992).


In der Fortsetzung interessieren die Leistungen der Deutschschweizer Siebtklässler im internationalen Vergleich und die Beziehung auf die vom Experimentiertest erfassten (a) materialen und (b) formalen Aspekte des handelnden Problemlösens. Anschliessend (c) wird am Beispiel der Experimentieraufgabe 'Tablette' aufgezeigt, wie sich die Lösungen von Experimentieraufgaben zur differenzierteren Bestimmung des Lernstandes nutzen lassen.
Ergebnisse

Die hier dargestellten Ergebnisse beruhen auf den Punktwerten für die Lösungsqualität. Aufgeführt wird mehrheitlich die durchschnittliche Lösungsqualität in Prozent der maximal möglichen Punktzahl pro Fach, Experimentieraufgabe oder Leistungserwartung. Ausgewählte Angaben zur Spezifität der Schülerantworten werden in die Beschreibung der Ergebnisse eingeflochten.

Fachleistungen

<table>
<thead>
<tr>
<th>Land</th>
<th>Schuljahr</th>
<th>Alter</th>
<th>Mathematik</th>
<th>Naturwissenschaften</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singapur</td>
<td>8</td>
<td>14.5</td>
<td>70 (1.7)</td>
<td>72 (1.8)</td>
</tr>
<tr>
<td>Schweiz</td>
<td>7</td>
<td>14.1</td>
<td>66 (1.5)</td>
<td>65 (1.0)</td>
</tr>
<tr>
<td>Schweden</td>
<td>7</td>
<td>13.9</td>
<td>65 (1.3)</td>
<td>63 (1.5)</td>
</tr>
<tr>
<td>Schottland</td>
<td>9</td>
<td>13.7</td>
<td>61 (2.2)</td>
<td>64 (1.5)</td>
</tr>
<tr>
<td>Norwegen</td>
<td>7</td>
<td>13.9</td>
<td>65 (1.1)</td>
<td>58 (0.8)</td>
</tr>
<tr>
<td>Tschechien</td>
<td>8</td>
<td>14.4</td>
<td>62 (1.7)</td>
<td>60 (1.3)</td>
</tr>
<tr>
<td>Kanada</td>
<td>8</td>
<td>14.1</td>
<td>62 (1.3)</td>
<td>59 (1.3)</td>
</tr>
<tr>
<td>Neuseeland</td>
<td>8.5 - 9.5</td>
<td>14.0</td>
<td>62 (1.3)</td>
<td>58 (1.5)</td>
</tr>
<tr>
<td>Spanien</td>
<td>8</td>
<td>14.7</td>
<td>52 (1.1)</td>
<td>56 (1.0)</td>
</tr>
<tr>
<td>Iran</td>
<td>8</td>
<td>14.6</td>
<td>54 (1.7)</td>
<td>50 (2.8)</td>
</tr>
<tr>
<td>Portugal</td>
<td>8</td>
<td>14.6</td>
<td>48 (1.3)</td>
<td>47 (1.2)</td>
</tr>
<tr>
<td>Zypern</td>
<td>8</td>
<td>13.8</td>
<td>44 (1.2)</td>
<td>49 (1.0)</td>
</tr>
<tr>
<td>Australien</td>
<td>8 oder 9</td>
<td>14.3</td>
<td>66 (1.5)</td>
<td>63 (1.1)</td>
</tr>
<tr>
<td>England</td>
<td>9</td>
<td>14.0</td>
<td>64 (1.0)</td>
<td>71 (0.9)</td>
</tr>
<tr>
<td>Niederlande</td>
<td>8</td>
<td>14.3</td>
<td>62 (1.5)</td>
<td>58 (1.4)</td>
</tr>
<tr>
<td>USA</td>
<td>8</td>
<td>14.2</td>
<td>54 (1.4)</td>
<td>55 (1.4)</td>
</tr>
<tr>
<td>Kolumbien</td>
<td>8</td>
<td>15.8</td>
<td>37 (2.5)</td>
<td>42 (1.4)</td>
</tr>
<tr>
<td>Rumänien</td>
<td>8</td>
<td>14.6</td>
<td>66 (2.0)</td>
<td>57 (2.0)</td>
</tr>
<tr>
<td>Slowenien</td>
<td>8</td>
<td>14.7</td>
<td>64 (1.0)</td>
<td>58 (1.4)</td>
</tr>
</tbody>
</table>

Internationaler Durchschnitt | 59 (0.3) | 58 (0.4)

Tabelle 1: TIMSS-Experimentiertest: Prozentsatz der richtigen Lösungen in Mathematik und Naturwissenschaften.


Abbildung 2: TIMSS-Experimentieraufgaben - Inhalte und Leistungserwartungen.


Bei der Problemstellung Möbeltransport sind die schweizerischen und der internationale Prozentsatz identisch. Ursache dafür sind die vielen Falschlösungen bei jenen zwei der acht Teilaufgaben, bei denen Schülerinnen und Schüler Möbelmodelle und Möbelmasse realen Möbelstücken zuordnen mussten. Ob unsere Jugendlichen sich nur mit Möbelmassen nicht auskennen oder generell eine vage Vorstellung von Längenmassen haben, kann anhand der vorhandenen Daten nicht beurteilt werden.

Dass der Prozentsatz der Experimentieraufgabe Verpackungen nicht deutlicher über dem internationalen Durchschnitt liegt, ist wesentlich darauf zurückzuführen, dass viele der Deutschschweizer Siebtklässler bei der dritten Teilaufgabe eine Schachtel anstatt des verlangten Faltpapieres hergestellt haben. Möglicherweise haben sie die Anleitung ungenau gelesen, sich vom Material zum Basteln verleiten lassen oder einfach geprüft, ob ihr Faltpapier einer vier Bälle fassenden Schachtel entspricht.


Die Auswertung nach Schultypen zeigt, dass die Gymnasiasten und die Sekundarschüler bei 11 der 12 Experimentieraufgaben signifikant bessere Leistungen erbrachten als die Realschüler. Obwohl die Testwerte der Gymnasiasten durchwegs etwas höher sind als jene der Sekundarschüler, ist der Unterschied zwischen Gymnasiasten und Sekundarschülern nur bei der Problemstellung Gummiband überzufällig.

Formale Kompetenzen

Um die Testleistungen der Jugendlichen hinsichtlich der aufgabenübergreifenden Kompetenzen zu vergleichen, ordnete TIMSS jede Teilaufgabe einer Kategorie der Leistungserwartungen zu. Für die folgenden Auswertungen wurden die Teilaufgaben ausgewählter Kategorien zu fünf Oberkategorien zusammengezogen (Tab. 2, S. 38).

In Mathematik wurde zwischen Routineverfahren (13 Teilaufgaben) und Problemlösen (21 Teilaufgaben) unterschieden. Alle Länder einschliesslich der Deutschschweiz erzielten bei den Routineverfahren höhere Testwerte als beim...

In den Naturwissenschaften wurden drei Teilbereiche unterschieden: Erforschen (16 Teilaufgaben), Routineverfahren (7 Teilaufgaben) sowie Problemlosen und Anwenden von Fachwissen (12 Items). Die höchsten Werte hatten die Deutschen Siebtklässler beim Erforschen, die zweithöchsten bei den Routineverfahren und die tiefsten beim Problemloesen und Anwendungen von Fachwissen. Auch hier ist der Unterschied zu den Achtklässlern aus Singapur bei den Routineverfahren am grössten. Und wiederum zeigen sich beim Problemloesen, aber auch beim Erforschen positivere Abweichungen bezogen auf den internationalen Durchschnitt als bei den Routineverfahren.

Aus diesen Ergebnissen lässt sich vorsichtig schliessen, dass die Achtklässlern aus Singapur die Deutschen Siebtklässler im Experimentiertest unter anderem deshalb überrannt haben, weil sie ein Schuljahr länger Zeit hatten, sich Routineverfahren anzuzeichnen. Umgekehrt kann man sagen, dass die Deutschen Siebtklässler beim Problemloesen und beim Erforschen trotz des fehlenden Schuljahres und des etwas tieferen Durchschnittsalters mit dem Spitzenreiter Singapur gut mithalten können. Weiter wird deutlich, dass sich die Verstehens-und Problemlöseorientierung unserer Lehrpläne im internationalen Leistungsvergleich auszahlt.

**Der Lernstand bei der Experimentieraufgabe ’Tabletten’**


Die Experimentieraufgabe Tabletten stammt aus dem Physikunterricht (Abb. 2). Sie ist in fünf Teilaufgaben gegliedert. Die Schüler(innen) und Schüler hatten kaltes und heisses Wasser, mehrere Becher, sechs Brausetabletten, einen Thermometer und eine Uhr mit Sekundenzeiger. Sie müssen ein naturwissenschaftliches Experiment durchführen, um begründete Aussagen über die Wirkung der Wassertemperaturen auf die Geschwindigkeit zu machen, mit der sich Tabletten auflösen.

---

Tabelle 2: Prozentsatz der richtigen Lösungen nach Fach und Leistungserwartung.

<table>
<thead>
<tr>
<th>Land</th>
<th>Mathematik Routine-</th>
<th>Mathematik Problem-</th>
<th>Naturwissenschaften Routine-</th>
<th>Naturwissenschaften Problemlösen/Fachwissen Erforschen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>verfahren</td>
<td>lösen</td>
<td>verfahren</td>
<td>Problemlosen/Fachwissen Erforschen</td>
</tr>
<tr>
<td>Singapur</td>
<td>80 (1.3)</td>
<td>62 (2.3)</td>
<td>75 (1.8)</td>
<td>59 (3.0)</td>
</tr>
<tr>
<td>Schweiz</td>
<td>76 (1.8)</td>
<td>60 (1.8)</td>
<td>63 (1.4)</td>
<td>55 (1.6)</td>
</tr>
<tr>
<td>Schweden</td>
<td>73 (1.3)</td>
<td>60 (1.6)</td>
<td>59 (1.9)</td>
<td>56 (2.3)</td>
</tr>
<tr>
<td>Schottland</td>
<td>75 (1.7)</td>
<td>52 (2.3)</td>
<td>69 (1.8)</td>
<td>48 (2.1)</td>
</tr>
<tr>
<td>Norwegen</td>
<td>75 (1.2)</td>
<td>58 (1.3)</td>
<td>57 (1.2)</td>
<td>48 (1.6)</td>
</tr>
<tr>
<td>Tschechen</td>
<td>73 (1.6)</td>
<td>56 (1.7)</td>
<td>57 (2.0)</td>
<td>53 (2.2)</td>
</tr>
<tr>
<td>Kanada</td>
<td>74 (1.4)</td>
<td>54 (1.3)</td>
<td>64 (2.2)</td>
<td>50 (1.6)</td>
</tr>
<tr>
<td>Neuseeland</td>
<td>72 (1.1)</td>
<td>55 (1.6)</td>
<td>65 (2.1)</td>
<td>47 (1.6)</td>
</tr>
<tr>
<td>Spanien</td>
<td>66 (1.4)</td>
<td>46 (1.3)</td>
<td>45 (1.8)</td>
<td>39 (1.6)</td>
</tr>
<tr>
<td>Iran</td>
<td>61 (1.8)</td>
<td>49 (1.8)</td>
<td>53 (3.4)</td>
<td>61 (2.0)</td>
</tr>
<tr>
<td>Portugal</td>
<td>66 (1.2)</td>
<td>36 (1.6)</td>
<td>47 (1.4)</td>
<td>32 (1.8)</td>
</tr>
<tr>
<td>Zypern</td>
<td>58 (1.3)</td>
<td>38 (1.4)</td>
<td>48 (1.7)</td>
<td>37 (1.9)</td>
</tr>
<tr>
<td>Australien</td>
<td>75 (1.4)</td>
<td>61 (1.9)</td>
<td>67 (1.9)</td>
<td>54 (2.0)</td>
</tr>
<tr>
<td>England</td>
<td>77 (1.1)</td>
<td>54 (1.3)</td>
<td>77 (1.4)</td>
<td>49 (2.0)</td>
</tr>
<tr>
<td>Niederlande</td>
<td>77 (1.7)</td>
<td>54 (1.5)</td>
<td>63 (1.7)</td>
<td>39 (1.9)</td>
</tr>
<tr>
<td>USA</td>
<td>64 (1.6)</td>
<td>49 (1.4)</td>
<td>61 (2.2)</td>
<td>43 (1.5)</td>
</tr>
<tr>
<td>Kolumbien</td>
<td>49 (2.7)</td>
<td>30 (2.7)</td>
<td>35 (2.4)</td>
<td>32 (2.2)</td>
</tr>
<tr>
<td>Rumänien</td>
<td>74 (1.9)</td>
<td>60 (2.4)</td>
<td>53 (2.5)</td>
<td>48 (3.3)</td>
</tr>
<tr>
<td>Slowenien</td>
<td>72 (1.2)</td>
<td>57 (1.1)</td>
<td>60 (1.3)</td>
<td>48 (1.5)</td>
</tr>
</tbody>
</table>

Internationaler Durchschnitt 70 (8.4) 52 (6.4) 59 (6.4) 47 (5.5) 60 (6.4)

TABLETTEN

An dieser Station braucht Du:
- Heisses und kaltes Wasser
- Mehrere Becher
- Einige Tabletten
- Einen Löffel
- Eine Uhr mit einem Sekundenzeiger
- Ein Thermometer
- Ein 30 cm-Lineal

Lies ALLE Anweisungen genau durch!

Deine Aufgabe:
Untersuche, wie schnell sich Tabletten bei unterschiedlichen Temperaturen auflösen.

Folgendes sollst Du tun:
- Plane eine Untersuchung, um herauszufinden, welche Wirkungen unterschiedliche Temperaturen auf die Geschwindigkeit haben, mit der sich Tabletten auflösen.

1. Schreibe Deinen Plan hier auf. Dein Plan sollte enthalten:
   - was Du messen willst.
   - wie viele Messungen Du machen willst.
   - wie Du Deine Messungen in einer Tabelle darstellen willst.

Wie schnell sich Tabletten bei unterschiedlicher Temperatur auflösen.
Zwischen 4-6 Messungen will ich machen.
Ich werde es in einer Tabelle, wo jeweils die Testergebnisse untereinander dargestellt ist zeigen.

2. Führe Deine Tests mit den Tabletten durch. Zeichne eine Tabelle und schreibe alle Deine Messungen auf:

<table>
<thead>
<tr>
<th>Temperatur</th>
<th>Zeit der Auflösung</th>
</tr>
</thead>
<tbody>
<tr>
<td>35°C</td>
<td>1 min 1 sek</td>
</tr>
<tr>
<td>39°C</td>
<td>56 sek</td>
</tr>
<tr>
<td>33°C</td>
<td>51 sek</td>
</tr>
<tr>
<td>36°C</td>
<td>47 sek</td>
</tr>
</tbody>
</table>

3. Welche Wirkung haben aufgrund Deiner Untersuchung verschiedene Temperaturen auf die Geschwindigkeit, mit der sich Tabletten auflösen?

Um so heißer das Wasser ist desto schneller lösen sich die Tabletten auf.


Beim warmen Wasser bewegen sich die kleinen Teilchen * schneller, damit kann sich alles besser auflösen und verteilten. Durch die ständige Bewegung der Moleküle, die viel Platz brauchen, wird die Tablette auflösen.

* (Molekül)

5. Wenn Du Deinen Plan ändern musstest, beschreibe Deine Änderungen und begründe, warum Du sie gemacht hast. Wenn Du Deinen Plan nicht ändern musstest, schreibe »Keine Änderung«.

Keine Änderung:
Leere den Inhalt Deiner Becher in das Waschbecken, trockne die Becher ab und hinterlasse alles so, wie Du es vorgefunden hast!

Abbildung 4: Testheft der Experimentieraufgabe 'Tablette' mit Schülerlösung.
Anmerkungen. Das Originaltestheft bestand aus drei A4 Seiten.

1. Bei der ersten Teilaufgabe müssen die Schülerinnen und Schüler einen Untersuchungsplan aufstellen.
Entscheidend ist, dass sie (i) die Untersuchung beschreiben und (ii) angeben, welche Variablen sie messen und (iii) welche sie konstant halten. Rund zwei Drittel der Schülerinnen und Schüler erfüllten diese Kriterien ganz oder teilweise (Abb. 5). Teil- und Falschslüsse entstanden dadurch, dass die Probanden ihr geplantes Vorgehen ungenau beschrieben, nur eine der beiden relevanten Größen (Wassertemperatur und Zeitdauer) nannten oder nicht angaben, dass sie mehr als einen Versuch vorsahen.

![Abbildung 5: Tablettenaufgabe - Lösungsqualität der Teilaufgaben.](image)

2. Anschliessend mussten die Schülerinnen und Schüler die Untersuchung durchführen.
drei Versuche protokolliert. Ob die betreffenden Probanden noch nicht wissen, dass Experimente mehrere Messungen erfordern, oder wie die Versuchsleiter annahmen, Hemmungen hatten, die Tabletten zu vergeuden, muss offen bleiben.

3. Welche Wirkungen haben aufgrund Deiner Untersuchung verschieden Temperaturen auf die Geschwindigkeit, mit der sich Tabletten auflösen?

Mit der Teilfrage drei wurde die Kompetenz der Schülerinnen und Schüler geprüft, Muster in Daten zu erkennen und Schlussfolgerungen zu ziehen. Beurteilt wurde, (i) ob die Schlussfolgerung mit den protokollierten Daten übereinstimmt und (ii) ob die Beziehung zwischen der Wassertemperatur und der Auflösedauer der Tablette beschrieben ist. Diese Teilaufgabe wurde am besten gelöst. Vereinzelt fehlten die Schlussfolgerungen.


Es wurde die Antwort erwartet, dass (i) mit steigender Wassertemperatur die kinetische Energie der Moleküle zunimmt und damit auch deren mittlere Geschwindigkeit, wodurch (ii) die Tablette heftiger attackiert und schneller zerrüttet wird. Unsere Siebtklässler waren mit dieser Frage überfordert, da der Fachbegriff ‘Kinetik’ in der Regel erst später unterrichtet wird. Die meisten wiederholten ihre Schlussfolgerung oder versuchten eine Erklärung aus ihrem Alltagswissen herzuleiten:

- Wenn es z.B. schneit und es 30° warm ware, wurde der Schnee gleich wieder schmelzen. … Auch beim Kochen ist es so … Mit kaltem Wasser kann man nicht kochen.
- Weil Flüssiges aufweicht oder auflöst.
- Wenn man Geschirr mit kaltem Wasser wascht, hat man mehr Muhe zum Trocknen als mit heissem Wasser.
- Ich denke, es ist wie bei der Schokolade, je tiefer die Temperatur, um so länger braucht auch sie, bis sie verlaufen ist. Die Kälte hat eine geringere Energie als die Wärme, sie braucht darum auch länger, bis etwas schmilzt oder vergeht.
- Wenn das Wasser wärmer ist, könnte es sein, dass sich der Sauerstoff, der sich in den Tabletten befindet, ausdehnt und deswegen die Tablettem schneller auflöst.
- Ich glaube, heisses Wasser ist aggressiver gegenüber Tabletten.
- Im Wasser hat es ganz viele kleine ‘Teilchen’, die mit zunehmender Wärme wärmer um sich irren (Ausdehnung des Wassers), so dass sie die Tabletten eigentlich auseinanderhängen.

5. Bei der fünften Teilfrage mussten die Schülerinnen und Schüler ihren Untersuchungsplan beurteilen.


**Diskussion**


**Der TIMSS-Experimentiertest gemessen an acht Gütekriterien**

nisse der Deutchschweizer Siebklässler erläutern und die Diskussion mit einigen Überlegungen zum Einsatz von Experimentieraufgaben im Unterricht abschliessen.

1. Die möglichen Wirkungen auf die Inhalte und die didaktische Gestaltung des Unterrichts müssen bedacht werden.

Dem Kriterium der Inhaltsäquivalenz entsprechend sind gute Experimentieraufgaben auf die in den Lehrplänen festgehaltenen formalen und materialen Ziele bezogen. Sie prüfen Inhalte, die in qualitativ hochstehendem, verstehensorientiertem Unterricht aufgebaut und geübt werden. Die Leistungsbewertung ist transparent, die Kriterien für die Beurteilung der Lösungen werden offengelegt.


(Lehrplan für die Volksschule des Kantons Zürich, 1991, S. 29).


2. Die Experimentieraufgaben müssen für alle Probanden gleichermassen fair sein.

Da Experimentieraufgaben kontextspezifisch sind, ist eine angemessene thematische Vielfalt unabdingbar, um die Voraussetzungen der Knaben und der Mädchen sowie Jugendlicher verschiedener Etnien ausgewogen zu berücksichtigen. Ferner müssen Kodiersysteme entwickelt werden, welche die intendierten Kompetenzen der Prüflinge erfassen.


Das Kodiersystem von TIMSS ermöglichte eine reliable und effiziente Beurteilung der Schülerleistungen. Es wurde deutlich, dass die präzise Formulierung der Beurteilungskriterien besonders bei Fragen nach Fachkonzepten entscheidend ist. Damit hier nicht Scheinwissen erfasst wird, darf nicht blass auf den Fachausdruck abgestellt werden, sondern es müssen die relevanten Beziehungen des Fachbegriffs operationalisiert werden.

3. Die limitierte Generalisierbarkeit der Testergebnisse muss berücksichtigt werden.


Ruiz-Primo und Shavelson (1996) zu folge braucht es zwischen 8 und 23 Experimentieraufgaben, um zuverlässige Angaben über den Leistungsstand eines Individuums zu machen. Bei Vergleichen auf Klassen- und Schulebene
kann man mit weniger Aufgaben auskommen. Im TIMSS-Experimentiertest
lässt ein Proband drei bis fünf Handlungsprobleme (mit je zwei bis acht Teil-
aufgaben). Insgesamt wurden 12 Experimentieraufgaben eingesetzt. Die von
TIMSS gewählte Anzahl Testaufgaben ist für Leistungsvergleiche auf der System-
ebene angemessen, sofern sie im Rahmen eines messtechnisch validen Designs
eingesetzt werden. Über die Leistungen einzelner Schüler oder Schülerinnen
lassen sich auf der Basis von drei bis fünf Experimentieraufgaben aber keine
Angaben machen.

4. Das Anspruchsniveau der Experimentieraufgaben muss bestimmt werden.

Um das Anspruchsniveau (cognitive complexity) einer Aufgabe zu bestimmen,
ist eine Aufgabenanalyse erforderlich, bei der das zur Aufgabenlösung erforder-
liche Sach- und Problemlösewissen spezifiziert und die Wechselwirkungen zwi-
sehen Aufgabe und Testnehmer berücksichtigt werden.

Bei TIMSS wurden Informationen über das Anspruchsniveau der Experimentieraufgaben durch Expertenurteile und einen Pilottest eingeholt. Zur
Beschreibung der erforderlichen Denk- und Problemlösekompetenzen wurde
das Konstrukt der Leistungserwartung verwendet. Die Kategorien wurden vor-
wiegend induktiv gewonnen. Dies mag mit ein Grund dafür sein, dass sich die bei TIMSS verwendeten Begriffe Routineverfahren und Problemlösen nicht
vollständig mit unserem kognitionspychologischen Verständnis decken.

5. Die Experimentieraufgaben müssen sich auf bedeutsame Inhalte des Faches
beziehen.

Es müssen mit Hilfe von Experten Inhalte gewählt werden, die für das entspre-
chende Fach relevant sind und einen hohen Bildungswert haben.

Diese Anforderung ist unseres Erachtens mit Bezug auf die formalen und die
materialen Aspekte des Experimentiertests erfüllt. Die Anwendungsprobleme
verlangen Routinen und Strategien, die für eigenständiges praktisches Problem-
lösen zentral sind. In materialer Hinsicht sind sie exemplarisch für grundlegen-
de mathematische und naturwissenschaftliche Sachverhalte (z.B. Blutkreislauf,
Wahrscheinlichkeit, Strahlensatz).

6. Die Experimentieraufgaben müssen den zu testenden Bereich gut
abdecken.

Im Gegensatz zu herkömmlichen Schulleistungstests, wo es vor allem darauf
ankommt, dass eine relevante Auswahl zentraler Fachkonzepte abgedeckt ist
(content sampling), wird bei Experimentiertests eher darauf geachtet, dass die
für wissenschaftliches Arbeiten wichtigen Problemlöse- und Experimentier-
tätigkeiten berücksichtigt werden (process sampling).

Dies trifft für den TIMSS-Experimentiertest zu. In formaler Hinsicht lösen
die Experimentieraufgaben ein reichhaltiges Problemlöse- und Experimentier-
verhalten aus. Hinsichtlich des Fachwissens ist das Spektrum dagegen sehr eng.
Fünf Physik- und eine Biologieaufgabe dürften den Bereich der Naturwissens-
chaften, der im schriftlichen TIMSS-Leistungstest die Physik, Biologie, Chemie,
Geographie und Umweltwissenschaften einschloss, kaum abdecken. Es ist
daher fraglich, ob die Ergebnisse der beiden Testteile in materialer Hinsicht ver-
gleichbar sind.

7. Die Experimentieraufgaben müssen für die Probanden interessant, bedeu-
tsam und lehrreich sein.

Eine gute Testaufgabe ist eine interessante und anschauliche Anwendungsauf-
gabe, lautet die Devise. Es ist bekannt, dass sich manche Jugendliche bei Lei-
stungstests nicht voll einsetzen, weil die Aufgaben unattraktiv sind. Auch bei
bei der TIMSS Erhebung war es für die Schülerinnen und Schüler vermutlich span-
nender, selbständig kleine Experimente durchzuführen als Fragen mit Mehr-
fach- und Kurzantworten zu bearbeiten, deren Lösungen keinen Einfluss auf die
Noten hatten.

8. Aufwand und Ertrag müssen in einem vertretbaren Verhältnis stehen.

Die Entwicklung und Durchführung von Experimentiertests ist verglichen mit
standardisierten Leistungstests sehr kosten-, personal- und zeitaufwendig. Der
Einsatz von Experimentiertests ist vorerst nur dann angezeigt, wenn Kompeten-
tzen gemessen und gefördert werden sollen, für die sich ökonomischere Testin-
strumente nicht eignen.

Zur Zeit erarbeiten in den USA verschiedene Forschergruppen Handreichungen
die Entwicklung ökonomischer und reliabler Experimentiertests (Shavelson,
et al. 1997). In England sind Experimentieraufgaben in Prüfungen und Unterr
richt etabliert (Gage & Berliner, 1992; Gee & Clackson, 1992). Im deutschespra-
chigen Raum hat dieses Testformat noch nicht Fuss gefasst. Dies mag damit
zusammenhängen, dass man hier keine Erfahrungen mit landesweiten Lei-
stungsvergleichen hat und somit kein Bedarf besteht, die unerwünschten
Nebenwirkungen zu kuriern. Solange die bislang angesprochenen messtechni-
schen Probleme vor allem mit Bezug auf die Generalisierbarkeit und Konstrukt-
validität (Leistungserwartungen) nicht befriedigend gelöst sind und für unseren
Sprachraum noch keine Wegeleitungen zur Konstruktion, Durchführung und
Auswertung vorliegen, können wir den Einsatz von Experimentieraufgaben zu
klassenübergreifenden Leistungsvergleichen nur mit Vorbehalt empfehlen.
Für den Unterricht sind sie unseres Erachtens vorzüglich geeignet.
Experimentieraufgaben im Unterricht

Ein Hauptziel unserer Lehrpläne ist der Erwerb eines flexiblen anwendbaren Sach- und Problemlöswissens. Experimentieraufgaben sind ein erfolgreiches didaktisches Arrangement, um Alltags- und Fachwissen zu vernetzen und positive Lernübertragung anzuregen. Allerdings muss gewährleistet werden, dass das Experimentieren nicht zum Manipulieren verkommt. Entscheidend ist, dass die Experimentieraufgaben als produktive Lernaufgaben in thematische Unterrichtseinheiten eingebettet werden und die Schülerinnen und Schüler erfahren, dass es beim Experimentieren nicht um spektakuläre Effekte geht, sondern darum (Schauble, Glaser, Duschl, Schulze, & John, 1995),

- die Beziehungen zwischen Ursachen und Wirkungen systematisch zu erforschen,
- ausgewählte Variablen begründet zu variieren,
- sinnvolle Muster zu entdecken,
- Vergleiche und Beurteilungen auf mehreren Dimensionen vorzunehmen und
- Beziehungen zwischen Experiment und Realität herzustellen.


Durch die Teilnahme am TIMSS-Experimentiertest haben wir eine interessante, prozessorientierte und anspruchsvolle Art der Leistungsmessung kennen gelernt, die unsere Lehrplänen entspricht, durch einen erweiterten kognitionspsychologischen Lernbegriff begründet ist und anderen Facetten der Intelligenz Rechnung trägt als viele herkömmliche Schulleistungstests. Aufgrund der noch nicht befriedigend gelösten, meist technischen Probleme und im Kontext unserer Lehr-Lern-und Prüfungskultur sehen wir den Nutzen von Experimentieraufgaben primär im unterrichtlichen Einsatz zur integrierten Förderung formaler und materialer Kompetenzen.

**Fussnoten**


2 Die Auswertung des schriftlichen Leistungstests auf der Sekundarstufe II liegt als Buch vor (Moser, Ramseier, Keller, & Huber, 1997). Die Ergebnisse für die Sekundarstufe II sind noch nicht veröffentlicht.

3 Die Schweiz ist das einzige Land im deutschen Sprachraum, das am TIMSS-Experimentiertest teilgenommen hat.

4 Die Bezeichnung 'Performance Assessment' wird uneinheitlich verwendet. Manche Leute verstehen darunter alle Formen der Leistungsmessung, die nicht auf dem multiple-choice Format beruhen.

5 Das Schlagwort 'higher order thinking' wird definitionsgemäß für kognitive Aktivitäten verwendet, beispielsweise nicht bloss abgerufen, sondern in irgendeiner Form transformiert werden müssen (Baker, 1990).

6 Um die Absurdität dieser Art von Berichterstattung zu verdeutlichen, schlagen Gage und Berliner (1992) vor, dass man einmal die Sterberate aller Spätaler eines Staates rangieren und der Leserschaft die Hintergrundinformationen vorenthalten könne.

7 Da bei der zweiten Teilaufgabe zwei Aspekte beurteilt wurden, enthält die Abbildung sechs Balken.

8 Die Erläuterung dieser Aussage würde den Rahmen dieses Textes sprengen.
Literatur


Des travaux pratiques pour développer l’ensemble des compétences des élèves. Résultats des tests de l’étude TIMSS

Résumé

Les expériences réalisées dans le cadre de l’étude TIMSS dans 44 classes de Suisse alémanique se sont appuyées sur des tests pratiques qui ont permis d’évaluer les prestations des élèves de 7e année en sciences et en mathématiques: il s’est agi pour eux de concevoir et de réaliser des expériences, d’en formuler les résultats, puis de dégager de ces données certaines tendances, pour en tirer des conclusions et les analyser à l’aide des outils conceptuels spécifiques de ces disciplines. Au niveau international, les résultats de ces élèves ont été remarquables, sans qu’il y ait à cet égard à distinguer entre filles et garçons. S’agissant maintenant d’une comparaison entre différentes classes, il existe encore certaines difficultés non résolues concernant les critères théoriques d’appréciation, et il convient donc de limiter pour l’instant la confiance que l’on peut placer dans ces tests d’évaluation comparative. L’introduction de ce type d’expériences pratiques visant à développer les aptitudes pratiques et théoriques des élèves en les mettant en situation de devoir résoudre un problème donné est entièrement compatible avec les grandes lignes de nos plans d’études.
Applicazioni pratiche per lo stimolo integrato di competenze formali e materiali - risultati dal test sperimentale TIMSS

Riassunto

Il test TIMSS è stato applicato in 44 classi della svizzera tedesca del VII anno scolastico (N=396) per problemi a carattere matematico e scientifico. I giovani dovevano pianificare e realizzare delle sperimentazioni, descrivere i risultati, dedurle i trend e spiegarli utilizzando la terminologia matematica e scientifica. Nel confronto internazionale gli allievi svizzeri hanno ottenuto risultati brillanti. Non si sono riscontrate differenze nel rendimento delle ragazze e dei ragazzi. Confronti tra le classi sono attualmente possibili solo con le dovute riserve a causa di problemi teorici di misurazione. L'attuazione dei compiti pratici per lo stimolo delle competenze formali e materiali nella risoluzione individuale di problemi e in un ambiente di insegnamento-apprendimento interattivo corrisponde agli obiettivi dei nostri piani di studio.

Practical Tasks for Integrated Furtherance of Formal and Material Competence - Results from the TIMSS-Experiment Test

Summary

For the TIMSS-experiment test, mathematical and scientific action problems were used to measure the achievement in 44 German-Swiss classes of the seventh year level. The students had to plan and carry out experiments, write notes of the results, gather tendencies, and draw conclusions and explain these using mathematical and scientific technical terms. In an international comparison, the German-Swiss seventh-formers come off very well. The achievements of the girls and the boys do not differentiate. Presently, comprehensive comparisons of achievement with experiment tasks cannot be recommended without reservations because of unsolved, measuring-theoretical problems. The use of practical tasks for integrated furtherance of formal and material competence in individual problem solving and in interactive teaching-learning environments corresponds to the central ideas of our curricula.
Praktische Anwendungsaufgaben zur integrierten Förderung formaler und materialer Kompetenzen

Erträge aus dem TIMSS-Experimentiertest

Rita Stebler, Kurt Reusser & Erich Ramseier