

## Die Heterogenität der Testergebnisse zu den Bildungsstandards Mathematik

Primärer Zweck des HarmoS Tests im Juni 07 war, die Aufgaben zu testen. Dabei sollten gleichzeitig verschiedene Fragestellungen beantwortet werden:

- Entsprechen die HarmoS – Items den Vorgaben des Rasch-Modells? Solche Items werden von leistungsstärkeren Lernenden besser bzw. häufiger gelöst als von weniger leistungsstarken Lernenden. Für die überwiegende Mehrheit der HarmoS Items trifft dies zu.
- Kann der HarmoS-Test bestätigen, dass die Kompetenz in Aspekten wie «Operieren und berechnen» oder «Explorieren und erforschen» zwar stark korreliert, dennoch trennscharf unterschieden werden kann. Mit einer Korrelation zwischen den verschiedenen Aspekten von jeweils etwa 0.8 kann auch diese Frage beantwortet werden.
- Gelingt es, genügend einfache Items zu validieren (bzw. empirisch eine genügend hohe Lösungshäufigkeit festzustellen) damit sie die auszuarbeitenden Mindeststandards ausreichend abstützen? Auch in diesem Punkt sind die Testergebnisse sehr zufrieden stellend.
- Wie stark ist das Lösungsverhalten sprachregional geprägt? Der Test ergab sprachregionale Unterschiede, die uns mehr als erstaunten. Dabei stand die Frage im Zentrum, welche Items von einer Region besonders gut bzw. besonders schlecht gelöst wurden.

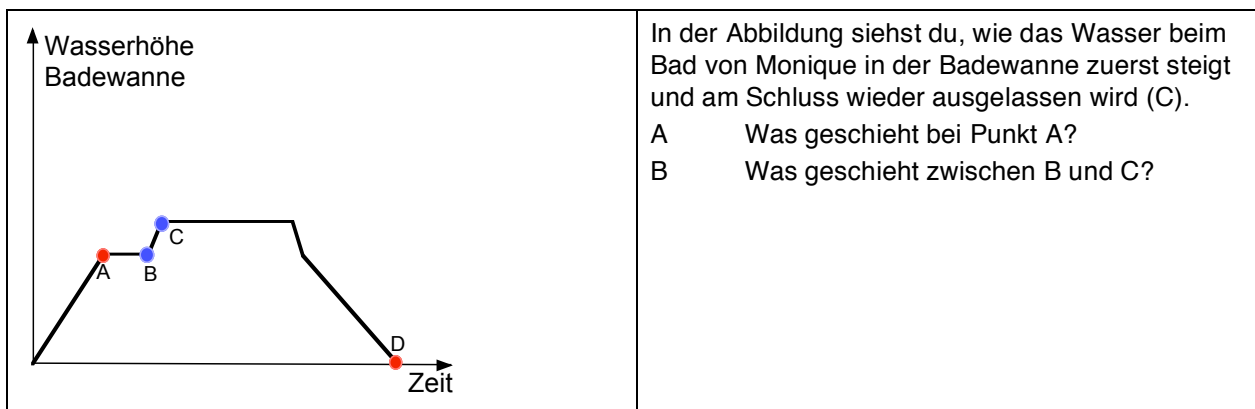
Zwei Beispiele zu Klasse 6 mögen dies verdeutlichen: M61407MO

Finde sich entsprechende Grössen und ordne sie einander durch Striche zu.

1 h 20 min	80 min
1800 sec	120 sec
2 min	eine halbe Stunde
1 $\frac{1}{2}$ h	900 sec
$\frac{1}{4}$ h	90 min

Diese Aufgabenstellung wurde in der deutschen Schweiz sehr gut gelöst (78%), während in den beiden andern Sprachregionen lediglich ca. 35% der Lernenden richtige Lösungen hervorbrachten.

Genau umgekehrt ist die Lösungshäufigkeit zu folgender Aufgabe: M61507FR



Etwa 20% der Lernenden der deutschen Schweiz lösten diese Aufgabe richtig, in den beiden andern Sprachregionen waren es jeweils etwa 50%.

Wer nun folgert, dass das Lösungsverhalten in der italienischen Schweiz und der Romandie ähnlich ist, irrt. Die Unterschiede  $dt - frz$ ,  $dt - ital$  und  $frz - ital$  sind jeweils etwa gleich gross.

Im Workshop wurden 12 Items mit einem auffälligen Sprachenbias analysiert und diskutiert. Dabei interessierten einerseits die Gründe, die zu diesen Ergebnissen führen konnten, andererseits wurden Vorschläge gemacht in welcher Art daraus Erkenntnisse für die Mathematik in der deutschen Schweiz genutzt werden können.