

Nora FELDT-CAESAR, Darmstadt

## **Möglichkeiten der Diagnose von Grundwissen und Grundkönnen durch ein adaptiv gestaltetes Testverfahren**

Um insbesondere an den Übergängen zwischen den Sekundarstufen und von der Sekundarstufe II zur Hochschule bei jeder Schülerin/jedem Schüler ein gewisses *Grundwissen und Grundkönnen* (Feldt, 2013) zu sichern, kommt einem dafür geeigneten Diagnoseinstrument besondere Bedeutung zu.

### **Anforderungen an ein Diagnoseinstrument**

Wenn ein Testinstrument das Grundwissen und Grundkönnen mehrerer Jahrgangsstufen umfassen soll, muss in der Regel eine Vielzahl relevanter Inhalte abgedeckt werden. Um dieser Inhaltsvielfalt in einem zeitlich notwendigerweise begrenzten Test gerecht zu werden, scheint es zunächst naheliegend, vorwiegend komplexe, d.h. mehrschrittige Items zu verwenden. Eine solche Verknüpfung von elementaren Inhalten und Handlungen wird zudem der Forderung nach einem im Sinne von Weinert *intelligenten* Grundwissen und Grundkönnen gerecht. Da Grundwissen und Grundkönnen die Voraussetzung für jedes weitere Lernen darstellt, ist es auf der anderen Seite besonders wichtig, eventuelle Defizite genau lokalisieren zu können. Elementare, einschrittige Aufgaben ermöglichen solch eine präzise Diagnose und erlauben damit, aus den Testergebnissen unmittelbar förderwirksame Rückschlüsse zu ziehen. Dies erscheint umso wichtiger, wenn ein digitales Testinstrument verwendet wird, das der Lehrkraft keinen Einblick in das Zustandekommen der Schülerlösungen bietet.

### **Elementarisierendes Testen als Lösungsansatz**

Ein im Sinne des *branched testings* (Kubinger, 2009) adaptiv gestaltetes Testverfahren bietet die Möglichkeit, sowohl komplexe als auch elementare Items je nach Bedarf zu kombinieren: Beim *elementarisierenden Testen* durchlaufen alle Schülerinnen und Schüler zunächst eine Hauptlinie von Testaufgaben, in der die Inhalte im Sinne der Zielformulierung von Grundwissen und Grundkönnen auch in verknüpfter Form getestet werden. Kann eine dieser Aufgaben nicht gelöst werden, wird man durch eine Schleife von Elementaritems geleitet, die jeweils einen für das Lösen der Hauptlinienaufgabe notwendigen *Elementarbaustein* (Feldt, 2013; Bruder et al., in Druck) fokussieren und somit zur Aufklärung des in der Hauptlinie aufgetretenen Fehlers beitragen. Nach dem Durchlaufen der elementarisierten Schleifenaufgaben wird mit der Bearbeitung der Hauptlinienaufga-

ben fortgefahren. Die Realisierung eines solchen Testverfahrens kann über ein digitales Testformat oder in diagnostischen Interviews erfolgen.

## Forschungsfragen und Erprobung

Zentral ist die Frage nach dem diagnostischen Mehrwert der vorgenommenen Elementarisierungen. Hier gilt es, für einzelne Aufgabenkomplexe (Hauptlinienaufgabe mit zugehörigen Elementaritems) Anhaltspunkte darüber zu gewinnen, inwieweit das Durchlaufen der elementarisierenden Schleife zur Fehleraufklärung beitragen kann. Eine weitere Forschungsfrage beschäftigt sich mit möglichen Lerneffekten beim Durchlaufen der elementarisierenden Schleifen. Diese Lerneffekte sind primär nicht intendiert, aufgrund des gezielten Ansprechens von Elementarbausteinen in den Schleifen aber durchaus zu erwarten. Zur Klärung dieser und weiterer Forschungsfragen wird derzeit ein Instrument zur Diagnose des am Ende der Sekundarstufe II verfügbaren Grundwissens und Grundkönnens im Bereich funktionaler Zusammenhänge mit SchülerInnen der Sekundarstufe II und StudienanfängerInnen im Rahmen des Mathematik-Vorkurses VEMINT und einer einführenden Mathematik-Lehrveranstaltung an der TU Darmstadt erprobt. Bei den StudienanfängerInnen wurden vier verschiedene Aufgabenkomplexe in verschiedenen Gruppierungen eingesetzt (Tangente: n=458; Extrema: n=459; Normale: n=332; Flächeninhalt: n=216). Zur Kontrolle des Lerneffekts wurde am Ende des Tests zu jeder der Hauptlinienaufgaben ein Parallelitem platziert, das im Falle einer ursprünglich nicht oder falsch gelösten Hauptlinienaufgabe angezeigt wurde.

## Diagnostisches Potential

	Gruppe 1 → (teilweise) Fehleraufklärung („Fehleraufklärungsquote“)	Gruppe 2 → keine Notwendigkeit einer Fehleraufklärung (Lerneffekt)	Gruppe 3 → keine Fehleraufklärung
Tangente (n=264)	62,0%	16,7%	21,3%
Extrema A (n=115)	51,3%	20,9%	27,8%
Normale (n=207)	78,3%	15,5%	6,3%
Flächeninhalt (n=104)	87,5%	1,9%	10,6%

*Die Angaben beziehen sich auf die Gruppe der TestteilnehmerInnen, die aufgrund eines Fehlers in der Hauptlinienaufgabe durch die elementarisierende Schleife geleitet wurde.*

Um Anhaltspunkte zum diagnostischen Mehrwert der Elementarisierungen zu erhalten, wurden die TestteilnehmerInnen, die die Hauptlinienaufgabe nicht lösen konnten und daher in die elementarisieren-

de Schleife geleitet wurden, in drei Gruppen eingeteilt: Gruppe 1 umfasst alle TestteilnehmerInnen, bei denen im Verlauf der Schleife mindestens ein Fehler registriert wurde, der auf mindestens einen fehlenden oder fehlerhaften Elementarbaustein schließen lässt. Der Anteil dieser Gruppe an allen TestteilnehmerInnen, die die Hauptlinienaufgabe nicht lösen konnten, wird

als „Fehleraufklärungsquote“ bezeichnet. Die Gruppen 2 und 3 umfassen hingegen die TestteilnehmerInnen, die bei allen elementarisierten Schleifenitems die richtige Lösung angeben konnten. Aus diagnostischer Sicht ist hier zunächst kein Mehrwert gegeben. Allerdings ist hier eine Unterscheidung sinnvoll, ob das Parallelitem im Anschluss an den Test gelöst werden konnte (Gruppe 2) der nicht (Gruppe 3). Kann die Anforderung des Hauptlinienitems schließlich doch bewältigt werden, so ist davon auszugehen, dass in irgendeiner Form ein Lernen stattgefunden hat. In diesem Fall ist es durchaus legitim, dass das Durchlaufen der elementarisierenden Schleife keinen Beitrag zur Fehlerrückmeldung leistet. Kritisch ist hingegen der Fall, der in Gruppe 3 eintritt: Hier findet weder ein Lernen noch eine Fehlerrückmeldung statt. Ein hoher Anteil von TestteilnehmerInnen in dieser Gruppe deutet darauf hin, dass eine zentrale Schwierigkeit der Hauptlinienaufgabe bei der Elementarisierung nicht erfasst wurde. (Ein geringer Teil der registrierten Fehler lässt sich vermutlich auch auf nicht-systematische Fehler zurückführen. Diese sind jedoch im verwendeten Testformat nicht kontrollierbar und werden daher bewusst ausgeblendet. Hierzu sollen anhand von zusätzlich erhobenen Interviewdaten weitere Anhaltspunkte gewonnen werden.) Die Anteile der jeweiligen Gruppen (vgl. Abb.) lassen darauf schließen, dass sich die eingesetzten Aufgabenkomplexe hinsichtlich der Qualität der Elementarisierung stark unterscheiden. Eine besonders hohe Fehlerrückmeldungsquote besitzt der Aufgabenkomplex „Flächeninhalt“, bei dem zusätzlich das diagnostische Potential von Distraktoren genutzt wird (Winter, 2011): Einer der zur Auswahl stehenden Distraktoren bildet den typischen Fehler ab, für eine unterhalb der x-Achse liegende Fläche statt des Flächeninhalts den orientierten Flächeninhalt anzugeben. Dieser Fehler macht 33,8% aller falschen Antworten aus. Da davon auszugehen ist, dass mit der Auswahl dieses Distraktors der Fehler vollständig aufgeklärt ist, wird der Testteilnehmer in diesem Fall nicht durch die elementarisierende Schleife, sondern direkt zum nächsten Hauptlinienitem geleitet. Werden beide Effekte zusammengenommen, erreicht man insgesamt eine Fehlerrückmeldungsquote von 91,7%. Der Anteil der kritischen Gruppe 3 reduziert sich gleichzeitig auf 7,0%. Dieses Beispiel zeigt, dass die Fehlerrückmeldungsquote durch den kombinierten Einsatz von elementarisierenden Schleifenaufgaben und Distraktoren mit diagnostischem Potential wesentlich verbessert werden kann.

### **Lerneffekte durch Reaktivieren**

Die Lösungshäufigkeiten der im Anschluss des Tests platzierten Parallelitems liegen bei etwa 30% bis 37%, sodass das Durchlaufen der elementarisierenden Schleifen zumindest zu kurzfristigen Lerneffekten zu führen

scheint. Lediglich beim Aufgabenkomplex „Flächeninhalt“ zeigt das Paralleliternur eine Lösungshäufigkeit von 16,3%. Der Grund hierfür liegt in dem erwähnten häufigen Fehler, der durch den Distraktor zwar aufgeklärt, jedoch nicht noch einmal gezielt in einer Schleifenaufgabe thematisiert wird, sodass hinsichtlich dieses Fehlers kein Lernanlass geboten wird. Bei den übrigen Aufgabenkomplexen zeigt die mittlere Anzahl der gelösten Schleifenitems einen signifikanten, gering bis mittelstark positiven Zusammenhang (Korrelationskoeffizient nach Pearson) mit dem Lösen des Paralleliterns. Einen Lerneffekt zeigen demnach vorwiegend die TestteilnehmerInnen, die viele Elementaritems lösen können. Das unterstützt die theoretisch auch zu erwartende Vermutung, dass die beobachteten Lerneffekte weniger auf ein Neulernen als vielmehr auf ein Reaktivieren von bereits erlernten, aber nicht mehr verfügbaren Kenntnissen, Fähigkeiten und Fertigkeiten zurückzuführen sind.

### **Zusammenfassung**

Die aktuellen Erprobungsergebnisse zeigen, dass mit dem elementarisierenden Testen ein Diagnoseverfahren entwickelt werden konnte, das die reguläre Teststruktur zunächst von der Einbindung vieler einzelner Elementaritems entlastet, um so verstärkt ein intelligentes Wissen und Können in den Blick zu nehmen. Gleichzeitig bleibt eine hohe diagnostische Aussagekraft erhalten, wobei die Fehleraufklärungsquoten zwischen den einzelnen Aufgabenkomplexen stark variieren. Insbesondere das Zusammenspiel von elementarisierenden Schleifen und diagnostischen Distraktoren scheint besonders wirksam. Neben dem eigentlichen Ziel der Gewinnung diagnostischer Informationen sind durch die vorgenommenen Elementarisierungen zumindest kurzfristige Lerneffekte zu beobachten, die vermutlich auf ein Reaktivieren von nicht mehr verfügbaren Kenntnissen, Fähigkeiten und Fertigkeiten zurückzuführen sind.

### **Literatur**

- Bruder, R., Feldt-Caesar, N., Pallack, A., Pinkernell, G. & Wynands, A. (in Druck). Mathematisches Grundwissen und Grundkönnen in der Sekundarstufe II. In W. Blum et al. (Hrsg.), *Bildungsstandards aktuell: Mathematik in der Sekundarstufe II*. Braunschweig: Schrödel.
- Feldt, N. (2013). Konkretisierung und Operationalisierung von Grundwissen und Grundkönnen durch ein theoriegeleitetes Vorgehen. In G. Greefrath et al. (Hrsg.), *Beiträge zum Mathematikunterricht 2013* (S. 308-311). Münster: WTM.
- Kubinger, K. D. (2009). *Psychologische Diagnostik, Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- Winter, K. (2011). *Entwicklung von Item-Distraktoren mit diagnostischem Potential zur individuellen Defizit- und Fehleranalyse*. Münster: WTM.