

Biologische Informationen aus unterschiedlichen Repräsentationen erschließen

Entwicklung und Überprüfung eines Testinstruments

Kathrin Klöpfel, Julia Schwanewedel & Jürgen Mayer

k.kloepfel@uni-kassel.de

Universität Kassel, Didaktik der Biologie,
Heinrich-Plett-Straße 40, 34132 Kassel

Zusammenfassung

Im vorliegenden Artikel wird die Konzeption und Erprobung eines Testinstruments zur Erfassung von Fähigkeiten von Schülerinnen und Schülern der Sekundarstufe I dargestellt. Mit Hilfe des Testinstruments soll eine zentrale (Teil-)Fähigkeit fachspezifischer Kommunikationskompetenz erfasst werden, nämlich das Erschließen biologischer Informationen aus unterschiedlichen Repräsentationen. Zur adäquaten Lösung des item-basierten Paper-Pencil-Tests müssen Schülerinnen und Schüler relevante Elemente in biologischen Texten und Bildern identifizieren, innerhalb der Repräsentation Zusammenhänge herstellen oder die Informationen mehrerer Repräsentationen integrieren. Das Instrument besteht aus 95 Items und wurde im Rahmen der hier vorgestellten Vorstudie in der 9. und 10. Jahrgangsstufe bei N = 328 Schülerinnen und Schülern eingesetzt. Ergebnisse zur Modell- und Itempassung werden im Rahmen des Artikels vorgestellt. In der Vorstudie haben sich 57 der insgesamt 95 eingesetzten Items (60 %) als raschhomogen und für die Erfassung der Schülerkompetenzen geeignet erwiesen. Die Befunde deuten darauf hin, dass die in der Itemkonstruktion zugrunde gelegte schwierigkeits erzeugende Dimension wie postuliert mit erhöhter empirischer Itemschwierigkeit korrespondiert.

Abstract

The present paper outlines the conception and subsequent trial of a paper-pencil test designed to assess secondary school students' abilities regarding one central (sub)capability of scientific communication competence: to process biological information from different representations. In order to properly engage in scientific communication, students have to identify relevant elements within biological texts and pictures, associate them with each other or integrate information of different representations. The task-based paper-pencil-test consists of 95 items and was distributed in 9th and 10th grade (N = 328). The presented

results focus particularly on aspects of model- and item-fit. In this test, 57 items (60 %) proved to be in accordance with the Rasch-model and adequate and effective to assess students' competencies. Findings indicate that the empirical item-difficulty correlates with the theoretically attributed item-difficulty.

1 Einleitung

Fachliche Kommunikationsfähigkeit wird als Teil naturwissenschaftlicher Grundbildung betrachtet (NORRIS & PHILLIPS, 2003). Sie soll im naturwissenschaftlichen Unterricht gefördert werden und wurde in Deutschland in den Bildungsstandards für die naturwissenschaftlichen Fächer im Kompetenzbereich Kommunikation verankert (KMK, 2005). Die Interpretation und Konstruktion von fachinhaltlich geprägten Repräsentationen z. B. Texten, Zeichnungen, Diagrammen und Tabellen, ist ein wichtiger Teil fachlicher Kommunikationsfähigkeit.

Repräsentationen spielen in den Naturwissenschaften in unterschiedlichen Bereichen eine Rolle. Für Naturwissenschaftler ist der Umgang mit Repräsentationen Teil des Arbeitsalltags und ein Mittel, um die eigenen Erkenntnisse in den Diskurs unter Fachkollegen einzubringen (GOLDMANN & BISANZ, 2002). Die nichtwissenschaftliche Öffentlichkeit erreichen neue Erkenntnisse meist über die Medien, in denen naturwissenschaftliche Sachverhalte durch populärwissenschaftliche Texte und Abbildungen veranschaulicht werden (GOLDMANN & BISANZ, 2002). Schülerinnen und Schüler sind im naturwissenschaftlichen Unterricht mit Schulbüchern, Arbeitsblättern und anderen Materialien konfrontiert, die Texte und Abbildungen enthalten. Die Fähigkeit relevante Informationen aus diesen zu erschließen, ist daher eine wichtige Ressource für das Lernen in naturwissenschaftlichen Fächern (YORE & TREAGUST, 2006). Die Fähigkeit mit fachinhaltlich geprägten Repräsentationen umgehen zu können, stellt für die Schülerinnen und Schüler somit eine Voraussetzung für das Verfolgen aktueller wissenschaftlicher Debatten und die spätere Teilhabe an der Gesellschaft als mündige Bürgerinnen und Bürger dar. Gleichzeitig ist sie eine Bedingung für das Lernen von Naturwissenschaften in der Schule.

Untersuchungen zu Fähigkeiten und Schwierigkeiten von Lernenden im Umgang mit Repräsentationen mit naturwissenschaftlichen Inhalten fokussieren häufig auf die Interpretation und Konstruktion von Repräsentationen sowie auf die Translation zwischen ihnen. COOK, WIEBE & CARTER (2008) untersuchten die Fähigkeiten von Highschool-Schülerinnen und -Schülern mit hohem und geringem Vorwissen beim Interpretieren von bildlichen Repräsentationen zur Osmose und Diffusion mit Hilfe von Eye-tracking und Interviews. Dabei

konnte festgestellt werden, dass Lernende mit hohem Vorwissen die für den fachlichen Zusammenhang relevanten Details einer Repräsentation identifizieren und Verbindungen zwischen den verschiedenen bildlichen Repräsentationen herstellen konnten (COOK, WIEBE & CARTER, 2008). Im Gegensatz dazu fokussierten Lernende mit geringem Vorwissen auf Oberflächenmerkmale und versuchten diese für die Integration der Repräsentationen zu nutzen (COOK, WIEBE & CARTER, 2008). Zu dem gleichen Schluss kommen KOZMA & RUSSELL (1997), die Chemikern bzw. fortgeschrittenen Chemiestudenten (Experten) und Schülerinnen und Schülern eines Colleges (Novizen) jeweils die gleiche Serie von Repräsentationen (Videoausschnitte, Diagramme, Animationen, Reaktionsgleichungen) zu einem chemischen Phänomen vorlegten. Bei der Aufgabe, die Repräsentationen zu clustern, zeigt sich die Orientierung an inhaltlichen bzw. Oberflächenkriterien der Experten bzw. Novizen ebenfalls (KOZMA & RUSSELL, 1997). Darüber hinaus waren die Experten deutlich besser in der Lage eine Repräsentation, wie eine chemische Reaktionsgleichung in eine andere beispielsweise in ein Diagramm umzuwandeln (KOZMA & RUSSELL, 1997). Mögliche Einflussfaktoren auf den Umgang mit Repräsentationen wurden von NITZ (2012) untersucht. Im Rahmen der Studie wurde dieser mithilfe von 15 Multiple-Choice-Items zu den Teilaspekten Interpretation, Konstruktion und Translation von Repräsentationen erfasst. In diesem Zusammenhang wird die Notwendigkeit weitere Aufgaben zu entwickeln, um das Konstrukt für den Biologieunterricht umfassend erfassen zu können, explizit betont.

Aus den beschriebenen Studien lassen sich Hinweise auf Fähigkeiten und Schwierigkeiten von Lernenden im Umgang mit Repräsentationen ableiten. In Bezug auf die Biologie steht eine systematische Erfassung der Fähigkeiten von Schülerinnen und Schülern im Umgang mit typischen Repräsentationen jedoch noch aus. Ziel der hier vorgestellten Studie ist daher die Entwicklung eines Testinstruments zur Erfassung der Fähigkeiten von Schülerinnen und Schülern der Sekundarstufe I im Bereich *Biologische Informationen aus unterschiedlichen Repräsentationen erschließen*. Das Instrument soll in der Lage sein einen wichtigen Teil fachlicher Kommunikationskompetenz zu erfassen, nämlich die Interpretation von verbal-sprachlichen und bildlichen Repräsentationen sowie die Integration von Informationen aus mehreren Repräsentationen.

2 Theorie

Die Bedeutung unterschiedlicher Repräsentationen für das Lehren und Lernen von Naturwissenschaften ist besonders international Gegenstand zahlreicher Studien. Es wird untersucht, inwieweit der Umgang mit multiple external representations (MERs) das Lernen fördern kann und beispielsweise zu einem vertieften Verständnis bei den Lernenden führt (AINSWORTH, PRAIN & TYTLER, 2011; TYTLER, PEARSON & PRAIN, 2006; WALDRIP, PRAIN & CAROLAN, 2010), aber auch inwieweit die Interpretation und Konstruktion von Repräsentationen sowie die Integration und Translation zwischen Repräsentationen den Lernenden Schwierigkeiten bereiten (COOK, WIEBE & CARTER, 2008; KOZMA & RUSSELL, 1997). MERs werden dabei als Kombinationen von mindestens zwei unterschiedlichen Typen von Repräsentationen betrachtet (u. a. WU & PUNTAMBEKAR, 2012).

2.1 Unterschiedliche Typen von Repräsentationen

Texte, Zeichnungen, Diagramme und Tabellen werden als Repräsentationen bezeichnet. Sie sind Objekte der Außenwelt und Beispiele für externe Repräsentationen. Im Rahmen des Verstehensprozesses konstruiert das verstehende Individuum eine interne, mentale Repräsentation (SCHNOTZ, 2003). In der vorliegenden Arbeit wird unter dem Begriff Repräsentation eine externe Repräsentation verstanden. Werden interne Repräsentationen thematisiert, wird der Begriff intern explizit als Zusatz verwendet.

Externe Repräsentationen lassen sich nach SCHNOTZ (2001) verbal-sprachliche (deskriptionale) und bildliche (depiktionale) Repräsentationen unterscheiden. Texte sind deskriptionale Repräsentationen, die einen Sachverhalt durch Symbole beschreiben. Dabei ist die Verknüpfung eines Symbols z. B. des Worts „Zelle“ mit dem zugehörigen Referenten durch eine Konvention festgelegt. Die Beziehungen der Sachverhalte werden durch explizite Relationszeichen z. B. Verben ausgedrückt (Schnotz, 2001). Im Bereich der depiktionalen Repräsentationen kann man zwischen realistischen und logischen Bildern unterscheiden. Realistische Bilder (z. B. Gemälde, Fotos und Strich- und Schemazeichnungen) zeichnen sich durch die Übereinstimmung von Struktureigenschaften der Repräsentation und des Sachverhaltes aus (Schnotz, 2001). Bei logischen Bildern stimmen die Relationen zwischen den Merkmalen innerhalb der Repräsentation und innerhalb des Sachverhaltes überein (Schnotz, 2001). Beispiele sind Strukturdiagramme, Flussdiagramme, Liniendiagramme und Tabellen. Die beschriebenen Repräsentationen wurden von von NITZ, NERDEL & PRECHTL (2012) für

die Biologie am Beispiel Fotosynthese spezifiziert. An dieser Stelle sollen Beispiele aus den Aufgaben der vorliegenden Studie angeführt werden (siehe Tab. 1).

Tabelle 1: Beispiele für verschiedene Arten von Repräsentationen in der vorliegenden Studie

Repräsentation	Aufgabentitel	Beschreibung der in der Aufgabe vorkommenden Repräsentationen
deskriptional / verbal- sprachlich	Energie- pyramiden	Ein Text über den Energiefluss im Ökosystem und die Energie auf verschiedenen Ernährungsstufen.
	Verdauung	Ein Text über die Rolle von Enzymen bei der Verdauung und den Abbau der Kohlenhydrate und der Eiweiße.
	Blut	Ein Text über die Bestandteile des Blutes und die Blutzellen.
	Transpiration	Ein Text über die Transpiration und die Abhängigkeit der Wasserabgabe von verschiedenen Umweltfaktoren bei Pflanzen.
depiktional / bildlich: realistische Bilder	Verdauung	Eine Schemazeichnung , die den Verdauungstrakt des Menschen zeigt.
	Blut	Eine rasterelektronen-mikroskopische Aufnahme der Blutzellen des Menschen.
depiktional / bildlich: logische Bilder	Energie- pyramiden	Eine Pyramidendarstellung , die die Energie auf verschiedenen Ernährungsstufen eines Ökosystems darstellt.
	Transpiration	Ein Liniendiagramm , das die Wasserabgabe eines Malvenblattes in Abhängigkeit von der Tageszeit zeigt.

2.2 Beschreibung und Operationalisierung der Fähigkeiten von Schülerinnen und Schülern im Bereich *Biologische Informationen aus unterschiedlichen Repräsentationen erschließen*

Der Umgang mit MERs erfordert ein gemeinsames Verstehen von Sprache und Bildern, wie es im Modell des integrierten Sprach-Bild-Verstehens beschrieben wird (SCHNOTZ & BANNERT, 1999; 2003). Im Modell wird davon ausgegangen, dass verbal-sprachliche und bildliche externe Repräsentationen auf gleiche Weise, jedoch in unterschiedlichen kognitiven Subsystemen verarbeitet werden. In einem ersten Schritt werden die Oberflächenmerkmale der externen Repräsentation subsemantisch verarbeitet. Die subsemantische Verarbeitung

beinhaltet Prozesse der Worterkennung und der syntaktischen Analyse bei verbal-sprachlichen externen Repräsentationen. Im Falle von bildlichen externen Repräsentationen umfasst die subsemantische Verarbeitung die visuelle Wahrnehmung des Bildes. Als Ergebnis der subsemantischen Verarbeitung entsteht eine interne, mentale Sprach- oder Bildoberflächenrepräsentation. Die subsemantische Verarbeitung verläuft ausschließlich datengeleitet und beinhaltet automatisierte Routinen. Damit ist sie vom Vorwissen und den Zielsetzungen des Individuums relativ unabhängig. Anschließend werden die internen, mentalen Oberflächenrepräsentationen des Textes und des Bildes im Rahmen der semantischen Verarbeitung in mehreren Schritten mithilfe von kognitiven Schemata analysiert und es wird jeweils ein mentales Modell konstruiert. Erst im dritten Schritt erfolgt die Integration der Informationen aus Text und Bild zu einem übergreifenden mentalen Modell (SCHNOTZ & BANNERT, 1999; 2003).

Im Sinne der hier angestrebten Erfassung der Fähigkeiten von Schülerinnen und Schülern im Bereich *Biologische Informationen aus unterschiedlichen Repräsentationen erschließen* genügt es nicht, die bei der Sprach-Bild-Verarbeitung individuell ablaufenden Prozesse zu beschreiben. Vielmehr sollen auf Basis der Prozesse Fähigkeiten beschrieben und differenziert werden, so dass sie in Form von Items operationalisiert und damit erfasst werden können. Eine solche Beschreibung in Anlehnung an das Modell von SCHNOTZ & BANNERT (1999; 2003) nimmt SEUFERT (2003) vor. Sie geht davon aus, dass Lernende, die mit einer Text-Bild-Kombination konfrontiert werden, zunächst die relevanten Informationen beider Repräsentationen identifizieren und untereinander verknüpfen müssen. Beide Prozesse laufen im Rahmen der von SCHNOTZ & BANNERT (1999, 2003) beschriebenen semantischen Verarbeitung für Texte und Bilder getrennt ab. Im Anschluss können die Informationen aus den unterschiedlichen Repräsentationen in Beziehung gesetzt werden. Dies entspricht dem Integrationsschritt im Anschluss an die semantische Verarbeitung bei SCHNOTZ & BANNERT (1999, 2003). Auf dieser Basis lassen sich für die vorliegende Arbeit die folgenden Teilfähigkeiten beschreiben:

Schülerinnen und Schüler müssen ...

... **innerhalb der Repräsentation Text relevante Elemente identifizieren (1a)** und zwischen diesen Elementen **Zusammenhänge herstellen (2a)** (SEUFERT, 2003).

... **innerhalb der Repräsentation Bild relevante Elemente identifizieren (1b)** sowie zwischen ihnen **Zusammenhänge herstellen (2b)** (Seufert, 2003).

... auf Basis der beiden zunächst genannten Schritte die **Informationen mehrerer Repräsentationen integrieren (3)** (Seufert, 2003).

In einem Schritt, den SEUFERT (2003) als *intra-representational coherence formation* bezeichnet, müssen die Lernenden Zusammenhänge zwischen den Repräsentationen herstellen und damit eine übergreifende Wissensstruktur des zugrunde liegenden biologischen Sachverhalts konstruieren.

Die Fähigkeiten der Schülerinnen und Schüler im Bereich *biologische Informationen aus unterschiedliche Repräsentationen erschließen* werden als latente, d. h. nicht direkt beobachtbare Variablen betrachtet. Diese Variablen werden in Form von Kompetenztestitems unterschiedlicher Schwierigkeit operationalisiert, sodass vom Abschneiden im Kompetenztest auf die zugrunde liegenden Fähigkeiten der Person geschlossen werden kann. Eine solche Art der Operationalisierung impliziert, dass die Items im ersten Schritt vor der Erfassung der Personenfähigkeiten hinsichtlich ihrer Eignung für die Kompetenzmessung erprobt werden. Auf Basis der statistischen Kennwerte findet eine Auswahl geeigneter Items statt. Diese erfolgte im Rahmen der vorliegenden Arbeit in einer Vorstudie, die im Weiteren vorgestellt wird.

3 Fragestellungen

Im Rahmen der Hauptstudie sollen die Fähigkeiten von Schülerinnen und Schülern der Sekundarstufe I im Bereich *biologische Informationen erschließen aus unterschiedlichen Repräsentationen* erfasst werden. Dies erfordert die Überführung des Konstrukts in ein Aufgabenkonstruktionsmodell, die Konstruktion eines geeigneten Testinstruments und seine Erprobung.

Überführung des Konstrukts in ein Aufgabenkonstruktionsmodell: **Welche Teilkonstrukte können beschrieben werden? Welche schwierigkeiterzeugende Dimension ist geeignet, um Aufgaben zu konstruieren?**

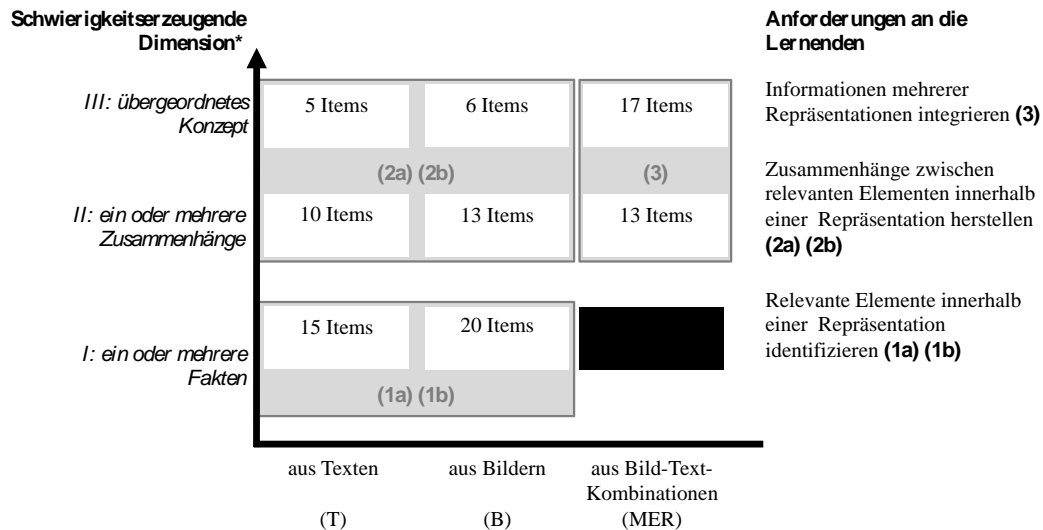
Itemkonstruktion: **Wie können Items zu unterschiedlichen Teilkonstrukten konstruiert werden? Wie können unterschiedlich schwierige Items konstruiert werden?**

Erprobung der Items in der Vorstudie: **Welche Items sind aufgrund der Itemkennwerte für eine spätere Kompetenzmessung geeignet?**

4 Methodik

Zur Erfassung der Fähigkeiten von Schülerinnen und Schülern im Bereich *Biologische Informationen aus unterschiedlichen Repräsentationen erschließen* wurde in der vorliegenden Studie ein Kompetenztest mit insgesamt 95 Items entwickelt. Diese Items wurden so konstruiert, dass sie nach dem FULL / NO

CREDIT System entweder mit richtig oder mit falsch kodiert werden. Bei der Itemkonstruktion wurde ein zweidimensionales Aufgabenkonstruktionsmodell zugrunde gelegt. Die Dimension *biologische Informationen erschließen* bildet die Inhalte ab, während die Dimension *Komplexität* als schwierigkeits erzeugend gilt (KAUERTZ et al., 2010) (Abb. 1). Bei der Itementwicklung werden die beiden Dimensionen gezielt miteinander in Beziehung gesetzt. Die so konstruierten Items lassen sich eindeutig auf beiden Dimensionen verorten.



* angelehnt an Kauertz (2008); Kauertz et al. (2010)

Abbildung 1: Aufgabenkonstruktionsmodell.

Alle Aufgaben enthalten einen Aufgabenstamm, der stets aus zwei Repräsentationen besteht. In der Beispielaufgabe „Verdauung“ besteht dieser aus einem Text und einem realistischen Bild (Abb. 2). Dieser Aufgabenstamm enthält das gesamte Fachwissen, das zur Lösung notwendig ist. Zu jedem Aufgabenstamm gehören mehrere Items. In den Items werden die Schülerinnen und Schüler aufgefordert Informationen aus einem Text, einem Bild oder einer Bild-Text-Kombination zu erschließen. Zur Graduierung wird die schwierigkeits erzeugende Dimension *Komplexität* verwendet. Auf dem Komplexitätsniveau I *ein oder mehrere Fakten* wird von Schülerinnen und Schülern lediglich die Berücksichtigung von einem oder mehreren unverbundenen Fakten erwartet. Auf Komplexitätsniveau II *ein oder mehrere Zusammenhänge* wird die Berücksichtigung einer oder mehrerer funktionaler Beziehungen erwartet. Auf Komplexitätsniveau III *übergeordnetes Konzept* müssen die in den Repräsentationen dargestellten Zusammenhänge generalisiert oder angewendet werden. Durch die Berücksichtigung der schwierigkeits erzeugenden Dimension soll gewährleistet werden, dass für alle drei Inhaltsdimensionen (T, B und MER) die Anzahl der Informationen, die auf einem Komplexitätsniveau, beispielsweise

se II zur Lösung des Items herangezogen werden müssen, vergleichbar ist. Die Testitems besitzen drei Antwortformate. Beim offenen Antwortformat müssen Schülerinnen und Schüler selbstständig Antworten aus einem Satz oder mehreren Sätzen generieren. Im Rahmen einer Kurzantwort müssen Antworten formuliert werden, die aus einem oder zwei Wörtern bestanden. Im geschlossenen Antwortformat (Multiple-choice, single-select) muss aus vier Antwortmöglichkeiten die richtige ausgewählt werden.

In Item fünf der Aufgabe „Verdauung“ (Abb. 2) müssen Schülerinnen und Schüler aus einem Text die enzymatischen Abbauschritte, die während der Verdauung ablaufen und die Organe in denen sie stattfinden entnehmen. Aus der Zeichnung müssen sie erschließen wie der Verdauungstrakt aufgebaut ist, sprich wie die Organe angeordnet sind. Um die enzymatischen Abbauschritte in die Reihenfolge zu bringen, in der sie im Körper ablaufen, müssen Schülerinnen und Schüler die Informationen aus beiden Repräsentationen anwenden. Das Item ist daher auf Komplexitätsstufe III bei Bild-Text-Kombination einzuordnen und stellt eine hohe Anforderung dar, da Informationen aus einer Bild-Text-Kombination integriert werden müssen. Im Sinne der Definition von MERs als aus mindestens zwei Repräsentationen bestehendes Gefüge werden solche Items als MER-Items bezeichnet.

Lena lernt für eine Klassenarbeit zum Thema Verdauung beim Menschen. Sie liest in ihrem Biologiebuch den Abschnitt über die Verdauung der Kohlenhydrate und der Eiweiße.

Zum Text gehört die folgende Abbildung:

Die Verdauung beim Menschen – der Abbau der Nährstoffe

Wichtig bei der Verdauung sind Enzyme. Sie spalten die Nährstoffe: Kohlenhydrate (Haushaltszucker und Stärke), Eiweiße und Fette.

Die Verdauung der Kohlenhydrate

Die Verdauung der Kohlenhydrate beginnt im *Mund*. Dort baut das Enzym Amylase einen Teil der mit der Nahrung aufgenommenen Stärke zu Malzzucker ab. Der Malzzucker wird im *Dünndarm* durch das Enzym Maltase in die einzelnen Bausteine zerlegt und in Blut und Lymphe aufgenommen.

Die Verdauung der Eiweiße

Das Enzym Pepsin baut die langen Eiweißketten aus dem Nahrungsbrei im *Magen* zu kürzeren Ketten ab. Im *Dünndarm* spaltet das Enzym Trypsin die kurzen Ketten in ihre einzelnen Bestandteile auf. Sie werden dort in Blut und Lymphe aufgenommen.

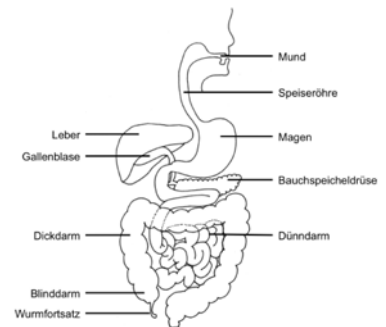


Abb. 1 Verdauung beim Menschen

Verdauung MER1505_Ver

Lena schreibt sich die einzelnen Abbauschritte für die Kohlenhydrate und die Eiweiße aus dem Text heraus. Sie möchte sie in die Reihenfolge bringen, in der sie auch im Körper ablaufen. Welche Reihenfolge muss Lena auswählen? Ergänze die Tabelle durch Ziffern.

Schritt	
	Das Enzym Amylase baut Stärke zu Malzzucker ab.
	Das Enzym Maltase zerlegt Malzzucker in die einzelnen Bausteine.
	Das Enzym Pepsin baut lange Eiweißketten zu kürzeren Ketten ab.
	Das Enzym Trypsin spaltet kurze Eiweißketten in einzelne Bestandteile auf.

Abbildung 2: Beispiel für einen Aufgabenstamm und ein Item.

Die Items wurden in 17 neunten und zehnten Klassen ($N = 328$) an verschiedenen Schulen in Hessen eingesetzt. 50,6 % der Probanden waren weiblich, 47,6 % männlich (1,8 % keine Angaben). In Bezug auf die verschiedenen Schulformen ergab sich die folgende Verteilung: 15,20 % Hauptschule, 27,20 % Gesamtschule, 57,60 % Gymnasium. Für den Einsatz im Rahmen der Vorstudie wurden 95 Items im Multimatrix-Design auf 23 Testhefte verteilt. Jedes Testheft enthielt zwischen 16 und 18 Items zu unterschiedlichen biologischen Themen. Bei der Zusammensetzung der Testhefte wurden die Items verschiedener Schwierigkeit sowie Items zu verschiedenen Repräsentationstypen systematisch verteilt. Nach einer kurzen Einführung standen den Schülerinnen und Schüler 35 Minuten für die Bearbeitung des Kompetenztests zur Verfügung. Die Auswertung erfolgte nach einem vorher festgelegten Kodierschema, das auf Basis der Schülerantworten ausgeschärft und um Beispiele für typisch richtige und falsche Antworten ergänzt wurde. Die Rohdaten der Testhefte wurden mit dem Programm Winsteps raschskaliert, sodass die Personenparameterschätzung zur Kompetenz der Probanden sowie die Schwierigkeit der Items auf einer Skala abgebildet werden konnten. Dabei wurde ein dichotomes Raschmodell zugrunde gelegt (BOONE & SCANTLEBURY, 2006). Im Rahmen der Auswertung der Vorstudien Daten wurden Aspekte der Modell- und Itempassung sowie Itemmerkmale untersucht. Aus diesem Grund wird im Rahmen der nachfolgenden Ergebnisdarstellung auf die Unterfrage **Welche Items sind aufgrund der Itemkennwerte für eine spätere Kompetenzmessung geeignet?** eingegangen. Die anderen Unterfragen beziehen sich auf die theoriebasierte Beschreibung der (Teil-)Fähigkeiten (vgl. Abschnitt 2.2) bzw. auf die Itemkonstruktion (vgl. Abschnitt 4).

5 Ergebnisse

Durch die Skalierung werden die 95 Items auf ihre Passung mit dem Raschmodell getestet. Für alle Items, die sich als raschhomogen erweisen, besteht ein Zusammenhang zwischen der Itemschwierigkeit und der Personenfähigkeit. Beide können daher auf einer Skala aufgetragen werden. Die Person-Item-Map stellt die geschätzten Personenparameter und die Itemschwierigkeit graphisch dar (Abb. 3). In der Bandbreite von -7 bis 6 sind links die Personen und rechts die Items dargestellt. Personen mit hohen Fähigkeiten und schwierige Items befinden sich in der Person-Item-Map oben, Personen mit geringen Fähigkeiten und einfache Items befinden sich in der Person-Item-Map unten. Liegen eine Person und ein Item auf einer Höhe, beträgt die Wahrscheinlichkeit, dass die Person das Item löst, 50 %. Die Person-Item-Map zeigt, dass es mit Hilfe der

eingesetzten Items gut gelingt ein breites Fähigkeitsspektrum abzudecken. Darüber hinaus weist der Gesamttest eine für die Stichprobe angemessene Schwierigkeit auf. Dies spiegelt sich in der Tatsache wider, dass die Mittelwerte der Personenfähigkeit und der Itemschwierigkeit auf einer Höhe bei 0 Logits liegen.

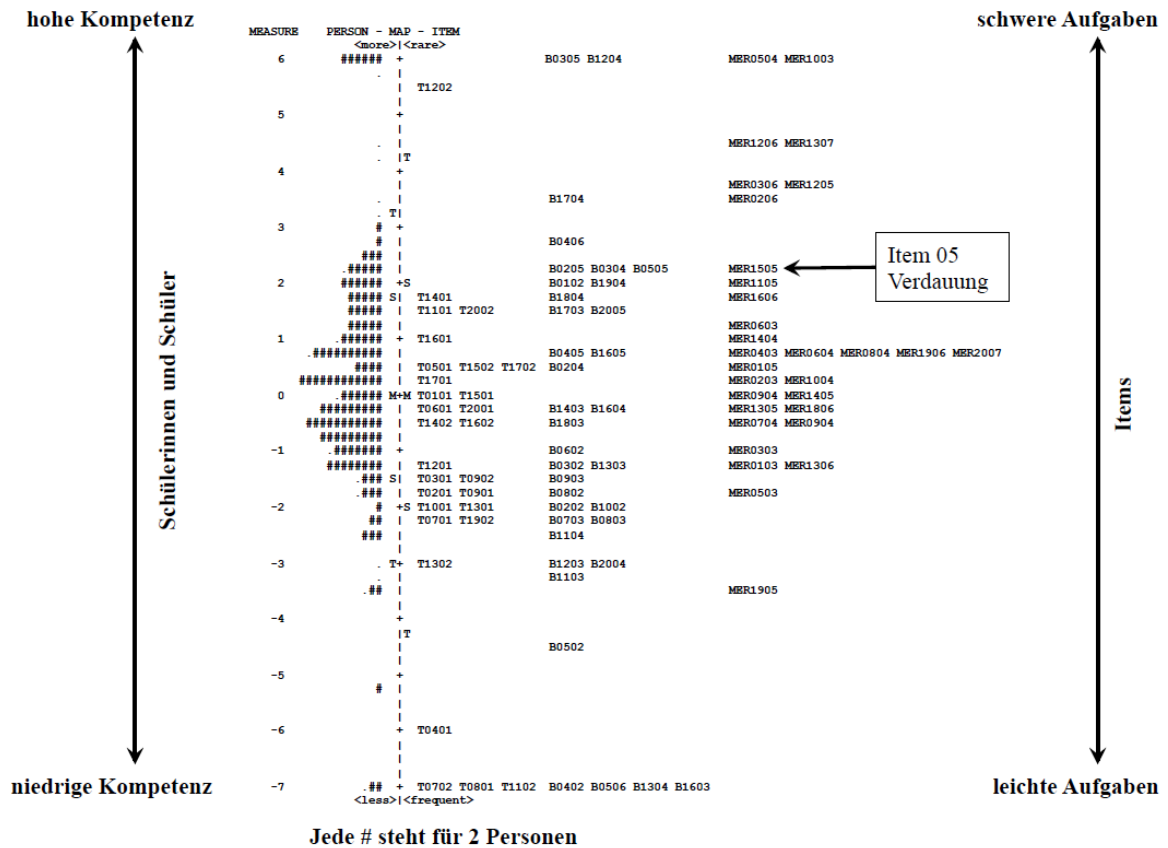


Abbildung 3: Person-Item-Map.

Im Rahmen der Itemanalyse wurden zunächst alle Items, die zu schwer oder zu leicht waren und daher einen Minimum- oder Maximum-Measure aufweisen, aus dem Itempool entfernt. Dies war für 9 Items (9,47 %) der Fall. Darüber hinaus wurde bei allen Multiple Choice Items eine Distraktorenprüfung durchgeführt. Dabei sollte kein Distraktor so unwahrscheinlich sein, dass er von weniger als 5 % der Probanden gewählt wurde. Aufgrund der Distraktorenprüfung wurden 10 Items (10,53 %) aus dem Itempool genommen. Als Kennwerte für die Beurteilung der Qualität der Items wurden zusätzlich der Item-Outfit und die Trennschärfe der Items betrachtet (LINACRE, 2012).

Der Item-Outfit hilft bei der Beurteilung der Items, indem er einen Anhaltspunkt für die Lösung der Items durch Personen an beiden Enden der Skala, also mit sehr hohen oder sehr geringen Fähigkeiten gibt. Beispielsweise sollte ein mittelschweres Item von Personen mit hohen Fähigkeiten gelöst und von Personen mit geringen Fähigkeiten nicht gelöst werden. Das von Winsteps ver-

wendete Item-Fit-Maß ist der MNSQ, der Item-Outfit entspricht dem unweighted MNSQ. Als produktiv für die Messung werden dabei Items mit einem MNSQ zwischen .5 und 1.5 angesehen (WRIGHT & LINACRE, 1994). Die Trennschärfe ist ein Maß, dafür, inwieweit ein Item zwischen guten und schlechten Probanden trennt. Die Trennschärfe in Winsteps entspricht nicht der klassischen Trennschärfe, da Winsteps für alle Items während der Berechnung eine Trennschärfe von 1.0 annimmt. Post-hoc wird die empirische Itemtrennschärfe geschätzt, wobei ein Wert geringer als 1.0 ein Anzeichen dafür ist, dass das Item schlechter zwischen guten und schlechten Probanden trennt als dies für ein Item dieser Schwierigkeit angenommen wird. Ein Wert höher als 1.0 zeigt an, dass das Item besser zwischen guten und schlechten Probanden trennt als dies für ein Item dieser Schwierigkeit anzunehmen ist. Für eine produktive Messung sollte die Trennschärfe zwischen .5 und 2.0 liegen (WRIGHT & LINACRE, 1994). Das in Abbildung 3 markierte Item der Beispielaufgabe „Verdauung“ besitzt beispielsweise einen MNSQ-Wert von .65 und eine Trennschärfe von 1.13. Das Item verhält sich demnach empirisch gemäß der Einstufung im Modell und ist ausreichend gut in der Lage zwischen fähigen und weniger fähigen Probanden zu unterscheiden. Insgesamt wurden 19 Items (20 %) aufgrund ihrer Kennwerte aus dem Itempool entfernt. 57 Items (60 %) konnten insgesamt mit angemessener Schwierigkeit nach positiver Distraktorenprüfung (im Falle der Multiple Choice Items) und aufgrund von guten Kennwerten in die Hauptstudie übernommen werden.

Auf Basis der Vorstudien­daten können vorläufige Aussagen bezüglich der Verteilung der Itemparameter (empirische Aufgabenschwierigkeit) differenziert nach Einordnung in die schwierigkeits­erzeugende Dimension gemacht werden. Betrachtet man die theoretisch zugeordnete Schwierigkeitsstufe mithilfe von Boxplots, korrespondiert eine höhere theoretisch zugeordnete Schwierigkeitsstufe des Items erwartungsgemäß mit einer höheren empirischen Itemschwierigkeit (Abb. 4).

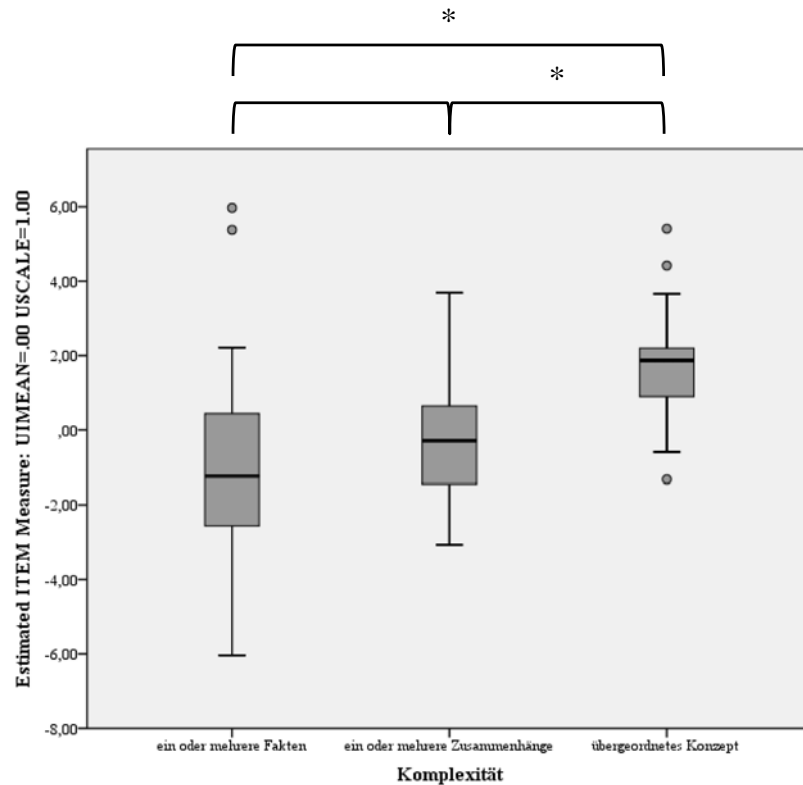


Abbildung 4: Verteilung der Itemparameter nach schwierigkeiterzeugender Dimension.

Demnach scheinen Items, in denen Schülerinnen und Schüler einen oder mehrere unverbundene Fakten berücksichtigen müssen, einfacher zu sein als solche, in denen sie Zusammenhänge, d. h. funktionale Beziehungen berücksichtigen müssen. Am schwierigsten scheinen Items zu sein, in denen die in der oder in den Repräsentationen dargestellten Informationen generalisiert oder angewendet werden sollen. Allerdings erweisen sich nur die Unterschiede zwischen der niedrigsten (ein oder mehrere Fakten) und der höchsten Komplexität (übergeordnetes Konzept) sowie der mittleren (ein oder mehrere Zusammenhänge) und der höchsten Komplexität als signifikant ($F(2,84) = 13.684$, $p < .05$; $r = .49$). Die Effektstärke deutet darauf hin, dass sich ein größerer Anteil der Varianz in den Daten über die schwierigkeiterzeugende Dimension aufklären lässt.

6 Diskussion und Ausblick

Basierend auf verschiedenen Arbeiten aus der fachdidaktischen Forschung und der Psychologie konnten die Fähigkeiten von Schülerinnen und Schülern im Umgang mit Repräsentationen beschrieben und für den Teilbereich *Biologisches Informationen aus unterschiedlichen Repräsentationen erschließen* ausdifferenziert werden. Auf Basis der inhaltlichen Ausdifferenzierung wurden

Items entwickelt, bei denen Schülerinnen und Schüler relevante Elemente in biologischen Texten und Bildern identifizieren, innerhalb der Repräsentation Zusammenhänge herstellen oder die Informationen mehrerer Repräsentationen integrieren müssen. Der Einsatz der 95 entwickelten Items in der Vorstudie von $N = 328$ ergab, dass die Raschhomogenität von 57 Items (60 %) gegeben ist und die als schwierigkeiterzeugend angenommene Dimension mit einer höheren empirischen Itemschwierigkeit korrespondiert. Im Februar / März 2013 wurde die Hauptstudie an ca. 1000 Schülerinnen und Schülern 9. und 10. Jahrgangsstufen in Hessen mit dem überarbeiteten Testinstrument durchgeführt. Die Aufgaben wurden für die Hauptstudie so aufbereitet, dass in allen Fällen in denen zu einer Aufgabe kein MER-Item mehr vorhanden war oder das MER-Item in dem entsprechenden Testheft nicht mitlief die Bilder in den Itemstamm der Bild-Items verlegt wurden. Dies eröffnet die Möglichkeit zu prüfen ob die jeweils andere Repräsentation das Informationen erschließen aus einem Text bei einem Text-Item bzw. das Informationen erschließen aus einem Bild bei einem Bild-Item beeinflusst. Ist dies nicht der Fall sollten die Text- und Bild-Items, die mit dem gesamten Aufgabenstamm mit beiden Repräsentationen laufen, die gleiche Schwierigkeit aufweisen wie diejenigen, die in der Variante laufen in der die Bilder in den Itemstamm der Bild-Items verlegt wurden. Auf Basis der Hauptstudien Daten werden weitergehende Analysen vorgenommen. Es wird beispielsweise untersucht, welche Kompetenzen Schülerinnen und Schüler im Bereich *Biologische Informationen aus unterschiedlichen Repräsentationen erschließen* besitzen und ob sich die beschriebenen Teilkonstrukte empirisch trennen lassen. Zusätzlich wird geprüft, wie hoch der Einfluss der Note in Biologie sowie weiteren Fächern (Deutsch, Mathematik) und der Lesekompetenz auf die operationalisierten Fähigkeiten ist. Darüber hinaus werden Strategien und Vorgehensweisen der Schülerinnen und Schüler im Umgang mit Repräsentationen mit einem Fragebogen erhoben und ihr Einfluss auf die Fähigkeit untersucht. Die Fähigkeiten von und die Anforderungen an Lernende in diesem Bereich fachlicher Kommunikationsfähigkeit werden somit im Rahmen der Arbeit ausdifferenziert. Zeigen sich Zusammenhänge zwischen Strategien und Fähigkeiten, können konkrete Ansatzpunkte für die Förderung im Biologieunterricht abgeleitet werden.

Zitierte Literatur

- AINSWORTH, S., PRAIN, V. & TYTLER, R. (2011): Learning Through Drawing in Science. *Science*, 333, 1096-1097.
- Boone, W. J. & Scantlebury, K. (2006): The Role of Rasch Analysis When Conducting Science Education Research Utilizing Multiple-Choice-Tests. *Science Education*, 90, 253-269.
- COOK, M., WIEBE, E. N. & CARTER, G. (2008): The Influence of Prior Knowledge on Viewing Graphs With Macroscopic and Molecular Representations. *Science Education*, 92, 848-867.
- GOLDMANN, S. R. & BISANZ, G. L. (2002): Toward a Functional Analysis of Scientific Genres: Implications for Understanding and Learning Processes. In: OTERO, J., LEÓN, J. A. & GRAESSER, A. C. [Hrsg.]: *The Psychology of Science Text Comprehension*. Mahwah, Lawrence Erlbaum, 19-50.
- KAUERTZ, A. (2008): Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben. Berlin, Logos.
- KAUERTZ, A., FISCHER, H. E., MAYER, J., SUMFLETH, E. & WALPUSKI, M. (2010): Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 135-153.
- KMK (2005): *Bildungsstandards im Fach Biologie für den mittleren Schulabschluss*. Neuwied, Luchterhand.
- KOZMA, R. B. & RUSSELL, J. (1997): Multimedia and Understanding: Expert and Novice Responses to Different Representations of Chemical Phenomena. *Journal of Research in Science Teaching*, 34(9), 949-968.
- LINACRE, J. M. (2012): WINSTEPS (Version 3.75.0). abgerufen auf www.winsteps.com
- NITZ, S. (2012): *Fachsprache im Biologieunterricht: Eine Untersuchung zu Bedingungsfaktoren und Auswirkungen*. Kiel, Christian-Albrechts-Universität.
- NITZ, S., NERDEL, C. & PRECHTL, H. (2012): Entwicklung eines Erhebungsinstruments zur Erfassung der Verwendung von Fachsprache im Biologieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 117-139.
- NORRIS, S. P. & PHILLIPS, L. M. (2003): How Literacy in Its Fundamental Sense Is Central to Scientific Literacy. *Science Education*, 87(2), 224-240.
- SCHNOTZ, W. (2003): Bild- und Sprachverarbeitung aus psychologischer Sicht. In: SACHS-HOMBACH, K. [Hrsg.]: *Was ist Bildkompetenz? Studien zur Bildwissenschaft*. Wiesbaden: Deutscher Universitäts-Verlag, 25-42.
- SCHNOTZ, W. (2001): Wissenserwerb mit Multimedia. *Unterrichtswissenschaft*, 29(4), 292-318.
- SCHNOTZ, W. & BANNERT, M. (2003): Construction and Interference in Learning from Multiple Representations. *Learning and Instruction*, 13, 141-156.
- SCHNOTZ, W. & BANNERT, M. (1999): Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Bild- und Textverstehen. *Zeitschrift für experimentelle Psychologie*, 46, 216-235.
- SEUFERT, T. (2003): Supporting Coherence Formation in Learning from Multiple Representations. *Learning and Instruction*, 13(2), 227-237.
- TYTLER, R., PEARSON, S. & PRAIN, V. (2006): Picturing Evaporation: Learning Science Literacy Through a Particle Representation. *Teaching Science*, 52, 12-17.
- WALDRIP, B., PRAIN, V. & CAROLAN, J. (2010): Using Multi-Modal Representations to Improve Learning in Junior Secondary Science. *Research in Science Education*, 40(1), 65-80.
- WRIGHT, B. D. & LINACRE, J. M. (1994): Reasonable Mean-Square Fit Values. *Rasch Measurement Transactions*, 8(3), 370.
- WU, H.-K. & PUNTAMBEKAR, S. (2012): Pedagogical Affordances of Multiple External Representations in Scientific Processes. *Journal of Science Education and Technology*, 21(6), 754-767.

YORE, L. D. & TREAGUST, D. F. (2006): Current Realities and Future Possibilities: Language and Science Literacy-Empowering Research and Informing Instruction. *International Journal of Science Education*, 28(2), 291-314.

