Evaluation der Aufnahmeprüfungen für die Berufsmittelschulen des Kantons Luzern

Institut für Lehren und Lernen

Prof. Dr. Werner Wicki
Roland Künzle

Unter Mitarbeit von Katharina Fischer, Hans-Peter Hodel, Reinhard Hölzl, Marco Hübner, Monika Mettler, Marianne Müller und Peter Widmer

Forschungsbericht Nr. 33
Pädagogische Hochschule Zentralschweiz
Hochschule Luzern

Evaluation der Aufnahmeprüfungen für die Berufsmittelschulen des Kantons Luzern

Werner Wicki
Roland Künzle

Unter Mitarbeit von Katharina Fischer, Hans-Peter Hodel Reinhard Hölzl, Marco Hübner, Monika Mettler, Marianne Müller und Peter Widmer

Bitte wie folgt zitieren:
Wicki, W. & Künzle, R. (2012) Evaluation der Aufnahmeprüfungen für die Berufsmittelschulen des Kantons Luzern. Luzern: Forschungsbericht Nr. 33 der Pädagogischen Hochschule Zentralschweiz, Hochschule Luzern.

Inhaltsverzeichnis

1. Ausgangslage, Ziele und Rahmenbedingungen 7

2. Evaluationsdesign und –methoden 9
2.1. Modul 1: Aufgabenanalysen 9
2.2. Modul 2: Vorhersage des BM-Abschlusses 9
2.2.1. Stichprobe und Datenanalyse 9
2.3. Modul 3: Vergleich der beiden Zulassungsvarianten 10
2.3.1. Stichprobe und Datenanalyse 11
2.4. Modul 4: Analyse der Kompetenzniveaus 11
2.4.1. Stichprobe und Datenanalyse 11
2.4.2. Methode zur Analyse der Kompetenzniveaus: 12

3. Aufgabenanalysen (Modul 1) 13
3.1. Beurteilung der Aufnahmeprüfungen Deutsch von 2009 und 2010 (Peter Widmer) 13
3.1.1. Übereinstimmung mit dem Lehrplan 13
3.1.2. Zielabdeckung mit dem Lehrplan 14
3.1.3. Prüfungsvergleich mit den Grundkompetenzen 19
3.1.4. Prüfungsvergleich mit dem Stoffplan 21
3.1.5. Schlusskommentar 22
3.1.6. Fazit 23
3.2. Beurteilung der Aufnahmeprüfungen Mathematik von 2009 und 2010 (Reinhard Hölzl und Marco Hübner) 24
3.2.1. Gewichtung (Punktevergabe) 24
3.2.2. Curriculare Validität 24
3.2.3. Kompetenzanalytische Beurteilung 25
3.3. Beurteilung der Aufnahmeprüfungen Englisch von 2009 und 2010 (Monika Mettler und Katharina Fischer) 30
3.3.1. Übereinstimmung mit dem Lehrplan 30
3.3.2. Globaler Eindruck 32
3.4. Beurteilung der Aufnahmeprüfungen Französisch von 2009 und 2010 (Hans-Peter Hodel) 32
3.4.1. Qualitätsmerkmale validen Prüfens 32
3.4.2. Lernziele /Niveaus Ende Sekstufe I gemäss Lehrplan und HarmoS 33
3.4.3. Angewandte Prüfmethoden 34
3.4.4. Analyse der Teilfertigkeiten der Prüfungen 35
3.4.5. Textproduktion 37
3.4.6. Analyse allgemeiner Eigenschaften der Prüfungen 38
3.4.7. Zusammenfassung 39

4. Vorhersage des BM-Abschlusses (Modul 2) 40

5. Vergleich der beiden Zulassungsvarianten (Modul 3) 41
5.1. Aufnahmeverfahren, Promovierungen und Dropout-Raten 41
5.2. Aufnahmeverfahren und Noten 42

6. Analyse der Kompetenzniveaus (Modul 4) 43
6.1. Raschskalierung AP Arithmetik und Algebra 2009 und 2010 43
6.2. Raschskalierung AP Deutsch 2009 und 2010 48

7. Fazit 51

8. Literatur 52

1. Ausgangslage, Ziele und Rahmenbedingungen

Die Aufnahme in eine Berufs- oder Fachmaturitätsschule erfolgt im Kanton Luzern entweder aufgrund einer Aufnahmeprüfung (AP) oder unter gewissen Bedingungen auch prüfungsfrei. Die AP wird von Sekundarschülerinnen und -schülern absolviert, die in der Regel im 6. Semester der Sekundarstufe Niveau A oder B besuchen. Es können sich aber auch Lernende aus C-Klassen anmelden.

Sekundarschüler/-innen, die das Niveau A besuchen und einen Notendurchschnitt von mind. 4.7 erreichen, werden prüfungsfrei aufgenommen.

Die Prüfung besteht in der vorliegenden Form seit ca. 10 Jahren, in den letzten Jahren wurde sie nur wenig verändert. Die Prüfungsaufgaben werden von Berufs- und Fachmittelschullehrpersonen und Lehrpersonen der Sekundarstufe I gemeinsam entwickelt.
Die Prüfung beinhaltet die drei Teile Deutsch, Mathematik und Fremdsprachen (Englisch und Französisch). Je nach Berufsmaturitätstyp werden (leicht) unterschiedliche Varianten vorgelegt und unterschiedliche Gewichtungen vorgenommen. Aufgrund der bisherigen Erfahrungen sind die Teilprüfungen in Mathematik und Französisch schwieriger und dadurch stärker selektiv als die andern Prüfungen.[1]

Die vorliegende Evaluation der Aufnahmeprüfung für die Berufsmaturitätsschulen (BM) geht auf eine Anregung der Berufs- und Fachmaturitätskommission des Kantons Luzern zurück. Sie interessierte sich einerseits dafür, inwieweit die Prüfungsaufgaben fair in Bezug auf das auf der Sekundarstufe vermittelte Wissen und Können sind (Blick zurück) und andererseits dafür, wie gut die Prüfung das Abschneiden am Ende der BM prognostizieren kann (Blick nach vorne).

Die Evaluation hatte zum Ziel, die Aufnahmeprüfungen der beiden Jahre 2009 und 2010 im Hinblick auf ihre inhaltliche Validität zu untersuchen. Darüber hinaus sollte untersucht werden, inwiefern das Prüfungsergebnis resp. die prüfungsfreie Aufnahme den erfolgreichen Besuch der Berufsmaturitätsschule vorhersagen kann (Vergleich der korrekt positiven mit den falsch positiven Aufnahmeentscheidungen).

Die Evaluation kann allerdings keine Aussage darüber machen, welchen Erfolg die aufgrund ihrer Prüfungsergebnisse abgewiesenen Kandidatinnen und Kandidaten gehabt hätten (korrekt negative vs. falsch negative Entscheidungen). Ein

[1] Die BM-Prüfungen 2010 und 2011 sind auf dem Internet unter dem folgenden Link publiziert: http://www.beruf.lu.ch/index/grundbildung/matur/berufsmatura/bm_aufnahmepruefung.htm (mittlerweile sind die BM-Prüfungen 2009 nicht mehr online verfügbar).
Nähere Informationen finden sich auf:
http://www.beruf.lu.ch/index/grundbildung/matur/berufsmatura/bm_aufnahmepruefung/aufnahmepruefung_informationen_09-7.pdf

weiterer Teil der Evaluation beschäftigt sich mit den Aufgabenschwierigkeiten der AP in den Fächern Mathematik und Deutsch.

Die Studie wurde vom Herbst 2011 bis Frühling 2012 durchgeführt. Ziel war, die Ergebnisse im Mai 2012 vorzulegen

2. Evaluationsdesign und –methoden

Diese Evaluation umfasst vier Module, die weitgehend unabhängig voneinander durchgeführt werden konnten.

2.1. Modul 1: Aufgabenanalysen

In diesem Modul wurden die Aufgaben der verschiedenen Teilprüfungen (Deutsch, Mathematik, Geometrie, Algebra, Englisch, Französisch) der Jahre 2009 und 2010 inhaltlich analysiert. Im Zentrum stand die Frage, ob die **Prüfungsaufgaben inhaltlich valide** sind und die zugrunde liegenden Lernziele mit denjenigen des Lehrplans der Sekundarstufe I (5. Semester) übereinstimmen. Diese Überprüfung erfolgte einerseits qualitativ und andererseits quantitativ in Form eines vierstufigen Ratings.
Zusätzlich wurden die Aufgaben im Hinblick auf die für deren Lösung notwendigen Kompetenzen analysiert.
Für die Durchführung der betreffenden Analysen konnten Fachdidaktiker/-innen der PHZ Luzern gewonnen werden.

2.2. Modul 2: Vorhersage des BM-Abschlusses

In diesem Modul wurde geprüft, ob und ggf. wie stark der Erfolg bei der Aufnahmeprüfung mit dem Erfolg bei der Abschlussprüfung zur Erlangung der Berufs- oder Fachmaturität zusammenhängt (prognostische Validität). Zu diesem Zweck wurden die **fachbezogenen Noten beider Prüfungen miteinander korreliert**.
Die Grundgesamtheit für diese Analyse bildete der **Jahrgang, der 2011 zum Abschluss gekommen** ist. Daraus wurde eine Stichprobe aus der Gruppe der Schüler/innen untersucht, von denen auch die Ergebnisse der Aufnahmeprüfung vorliegen. Neben den Korrelationen innerhalb der gleichen Fächer wurden auch Zusammenhänge zwischen verschiedenen Fächern im Quer- und Längsschnitt berechnet.
Die untersuchte Leitfrage bezog sich somit in diesem Modul auf die prognostische Validität der Aufnahmeprüfung.

2.2.1. Stichprobe und Datenanalyse

Zum Zwecke der vorliegenden Evaluation wurden – vor allem aus pragmatischen Gründen – drei Berufsmittelschulen des Kantons Luzern, die sich in der Stadt

Luzern befinden, ausgewählt: KBZ, BBZB, FMZ. Diese Schulen stellten dem Evaluationsteam die Noten der Aufnahme- und Abschlussprüfungen aller Schüler und Schülerinnen des berücksichtigten Jahrgangs (s.o.) in Form von Excel-Dateien zur Verfügung.[2]
Die analysierte Stichprobe umfasste ursprünglich 534 Schülerinnen und Schüler. Sie setzte sich innerhalb der drei Schulen wie in Tabelle 1 dargestellt aus lehrbegleitenden, schulorganisierten, berufsgleitenden und vollzeitlich organisierten Formen zusammen. Nach Abzug von Schülerinnen und Schülern, die vorzeitig ausgetreten waren, die Ausbildung gar nicht angetreten hatten oder (nach bestandener Aufnahmeprüfung) erst ein Jahr später (im SJ 11/12) eintraten, verblieben 462 Schülerinnen und Schüler in der Stichprobe.

Tabelle 1: Stichprobe der beteiligten Schulen

Schule	Mädchen	Jungen	Total
KBZ BMS Kaufm. Richtung lehrbegleitend	76	32	108
BBZ BMS technisch lehrbegleitend	4	45	49
BM kaufmännisch schulorganisiert	72	73	145
BMS 2 technisch Vollzeit	3	102	105
BMS 2 Gesundheit und Soziales	65	9	74
BMS 2 technisch berufsbegleitend	5	24	29
BMS 2 gestalterisch Vollzeit	12	12	24
Gesamt	237	297	534

Die Exceldateien drei Schulen wurden zusammengeführt und in das Statistikprogramm SPSS eingelesen. Mit dem SPSS wurden die statistischen Analysen durchgeführt.

2.3. Modul 3: Vergleich der beiden Zulassungsvarianten

In diesem Modul wurden die Schüler/-innen, die **prüfungsfrei** zugelassen worden waren, mit den Schüler/-innen **verglichen**, die die **Aufnahmeprüfung bestanden** hatten. Im Zentrum stand dabei die Frage, ob das Kriterium für den prüfungsfreien Zugang (Notendurchschnitt von mind. 4.7 für Sek-A) angemessenen ist oder nicht. Mögliche Indikatoren waren hier Schulnoten, provisorische Promovierungen und Drop outs. Die Stichprobe besteht aus Lernenden, die im Jahr 2010 zugelassen wurden. Grundlage für den Vergleich waren die Indikatoren ein Jahr nach dem Eintritt (d.h. Ende SJ 2011).

[2] Die bei den betreffenden Prüfungen erreichten Punkte lagen den Schulen nicht in elektronischer Form vor. Aus Gründen des Aufwands wurden deshalb für die Analysen dieses Moduls anstelle der (genaueren) Testpunkte die Noten verwendet.

2.3.1. Stichprobe und Datenanalyse

Wie in Modul 2 setzte sich die Stichprobe aus Lernenden von drei Berufsmaturitätsschulen auf dem Platz Luzern zusammen. Wir erhielten verwendbare Daten aus den fünf in Tabelle 2 aufgeführten Ausbildungsvarianten. Insgesamt umfasst diese Stichprobe die Angaben von N = 360 Schülern/innen. Nicht in dieser Stichprobe sind aus technischen Gründen die Daten der Varianten „BMS II technisch Vollzeit“ und „BMS II gestalterisch Vollzeit“.

Tabelle 2: Stichprobe und beteiligte Schulen (Modul 3)

Schule	Mädchen	Jungen	Total
KBZ BMS Kaufm. Richtung lehrbegleitend	72	29	101
BBZ BMS technisch lehrbegleitend	10	69	79
BM kaufmännisch schulorganisiert	34	58	92
BM2 Gesundheit und Soziales	44	9	53
BMS II technisch berufsbegleitend	9	26	35
Gesamt	169	191	360

Wiederum wurden Excel-Files der Schulen in eine SPSS-Datei übertragen.
Die Aufnahmevarianten wurden mithilfe von Chi-Quadrat-Tests bezüglich Drop-outs und Promovierungen miteinander verglichen. Ausserdem wurden die Aufnahmevarianten im Hinblick auf die erreichten Noten ein Jahr nach dem Eintritt miteinander verglichen (t-Tests, Mann-Whitney-Tests).

2.4. Modul 4: Analyse der Kompetenzniveaus

Das Ziel dieser Analyse bestand darin, die Prüfungsaufgaben der Fächer Arithmetik / Algebra und Deutsch im Hinblick auf die für deren Lösung erforderlichen Kompetenzen zu analysieren. Im Zentrum stand hier die Frage, ob die Prüfungsaufgaben innerhalb einer Prüfung verschiedenen Kompetenzniveaus zugeordnet werden können, ob also ein Mix unterschiedlich schwieriger Aufgaben vorgelegt wurde.
Diese Analyse ergänzt und erweitert die im Modul 1 vorgenommenen Einschätzungen (zur inhaltlichen Validität). Die Beschränkung auf zwei Fächer und zwei Jahrgänge (2009 und 2010) war rein pragmatisch-ökonomisch motiviert.

2.4.1. Stichprobe und Datenanalyse

Die Analyse der Kompetenzniveaus erfolgte anhand einer Stichprobe von je N = 200 Aufnahmeprüfungen in den Fächern Deutsch und Arithmetik/Algebra für die

Jahre 2009 und 2010 des FMZ (FMS, GL, GT, GV, WML)[3] verwendet. Die Prüfungsergebnisse und die Punktzahlen pro Aufgabe wurden vom Evaluationsteam elektronisch erfasst. Um die Zahl der Eingabefehler möglichst gering zu halten, wurde das Programm Epidata verwendet.[4]
Um die Analysen durchführen zu können, wurden nachträglich demographische Daten der Prüfungsteilnehmer (Geschlecht, Alter, Nationalität) beim FMZ in eingeholt und in die SPSS-Datei eingetragen.

2.4.2. Methode zur Analyse der Kompetenzniveaus

Um der Frage nachzugehen, wie gut man anhand der vorliegenden Prüfungen in Arithmetik / Algebra und Deutsch unterschiedliche Fähigkeiten einschätzen kann, wurde eine Raschskalierung durchgeführt. Bei dieser Skalierungsmethode (Skalierung = Messbarmachung) werden die Fähigkeiten von Personen und die Schwierigkeiten von Aufgaben miteinander verglichen. Personen mit mehr Fähigkeiten können auch die schwierigeren Aufgaben lösen. Durch Modellberechnungen werden die Aufgaben gemäss ihrer Schwierigkeit angeordnet und den Personenfähigkeiten gegenübergestellt. Für eine erfolgreiche Raschskalierung müssen gewisse Voraussetzungen (Eindimensionalität, Differential Item Functioning (DIF), lokale stochastische Unabhängigkeit) erfüllt sein (vgl. Bühner 2006, Strobl 2010).

[3] GL: Berufsmaturität G & S lehrbegleitend, GV: Berufsmaturität G & S Vollzeit, GT: Berufsmaturität G & S Teilzeit, WML: Wirtschaftsmittelschule, FMS: Fachmittelschule

[4] Da die Aufnahmeprüfungen in Algebra und Deutsch in ihrem Aufbau sehr verschieden waren, gestaltete sich auch die Dateneingabe unterschiedlich aufwändig. Bei den Algebraprüfungen wurde eine Eingabeschablone für das Punkteraster auf der Frontseite verwendet, da hier die Punktezahl pro Aufgabe klar ersichtlich ist. Für die Deutschprüfungen musste eine Eingabeschablone für die ganze Prüfung erstellt werden, da die Frontseite nur Punktesummen für die verschiedenen Prüfungsteile: "Textverständnis", "Wortschatz, Stil", "Grammatik, Interpunktion" und „Texte verfassen" enthält.

3. Aufgabenanalysen (Modul 1)

Die Aufgabenanalysen wurden von Fachdidaktikern der PHZ Luzern vorgenommen: Peter Widmer analysierte die Deutschprüfungen, Reinhard Hölzl und Marco Hübner die Mathematikprüfungen, Monika Mettler und Katharina Fischer die Englischprüfungen und Hans-Peter Hodel die Französischprüfungen.
Ihre Ausführungen wurden lediglich redaktionell bearbeitet.

3.1. Beurteilung der Aufnahmeprüfungen Deutsch von 2009 und 2010 (Peter Widmer)

Zur inhaltlichen Evaluation der Aufnahmeprüfung Deutsch wurde wie folgt vorgegangen: Die Aufgaben wurden gelöst und ein kurzer Kommentar zu den einzelnen Aufgabenstellungen formuliert. Dann folgten die Zuordnung der einzelnen Aufgabenstellungen zu den Lernzielen des Lehrplans und ein Rating zur inhaltlichen Übereinstimmung mit Lehrplan. Die Übereinstimmung mit dem Lehrplan wurde zudem qualitativ beschrieben. Darauf folgten Vergleiche der Aufgabenstellungen einerseits mit den Grundkompetenzen für die Schulsprache gemäss nationalen Bildungsstandards, andererseits mit dem Stoffplan für die Aufnahmeprüfung. Am Ende des Abschnitts wird ein Fazit gezogen.

3.1.1. Übereinstimmung mit dem Lehrplan

Die Beurteilung der Übereinstimmung mit dem Lehrplan basierte auf den folgenden Dokumenten:

- Lehrplan Deutsch Sekundarschule, Kommission Deutsch der Innerschweizer Erziehungsdirektoren 2. Auflage 2002
- Deutsch Sekundarstufe I: Übersicht Grobziele, Amt für Volksschulbildung des Kantons Luzern, Lehrplananpassungen 2006
- Lehrplan Deutsch: Klärung normativer Bereich für die Sekundarstufe Niveau C; B; A, Fachberatung Deutsch

Der Lehrplan Deutsch ist gegliedert in 7 Richtziele mit einer unterschiedlichen Anzahl Grobzielen und thematischen Bestimmungen.
Richtziele und Grobziele sind nicht als Kompetenzen beschrieben, sondern in Form von offenen Formulierungen. Die thematischen Bestimmungen geben Auskunft über nähere inhaltliche Angaben, machen aber keine Aussagen über genaue Ziele und Anforderungen.
Die Richtziele formulieren Einstellungen und Verhaltensweisen, die sich Lernende aneignen müssen in den Bereichen: Sprachliches erfassen, Ordnen, Denken;

Mündliche Kommunikation; Schriftliche Kommunikation; Mitteilungsvermögen; Medien und das sprachliche Kunstwerk.
Die Richtziele und die dazugehörenden Grobziele sind nicht nach den eigentlichen Bereichen des Deutschunterrichtes, Lesen, Sprechen / Hören, Schreiben, Sprachreflexion, geordnet.
Für den Bereich „Lesen“ gibt es kein explizites Richtziel, Grobziele zum Lesen müssen in anderen Richtzielen gesucht werden.
Entsprechend lassen sich die Prüfungsaufgaben nicht einzelnen Grobzielen zuordnen, sondern es sind verschiedene Grobziele, die mit den Aufgaben verbunden sind.
Deshalb ist eine genaue Beurteilung der Lehrplanübereinstimmung nicht einfach und entspricht weitgehend einer Interpretationsfrage.

3.1.2. Zielabdeckung mit dem Lehrplan

Die Aufnahmeprüfungen Deutsch 2009 und 2010 umfassen die Prüfungsteile: Textverständnis, Wortschatz / Stil, Grammatik / Interpunktion / Orthografie, und Texteschaffen.

In den folgenden Tabellen wurden die einzelnen Aufgabenstellungen kommentiert. Ausserdem wurden den einzelnen Aufgabenstellungen die Lernziele aus dem Lehrplan und ein Rating, 1 - 4 zugeordnet.

Tabelle 3: Aufnahmeprüfung 2009 Deutsch

Aufgabe	Max. Punkt-zahl	Kommentar zur Aufgabe durch Prüfungsbeurteiler	Übereinstimmung Lehrplanziel(e)	Rating (1-4) [5]
Textverständnis				
1. Meinung des Vaters	2	Klare Anweisung: Zwei Aspekte, eigene Worte, ganzer Satz	1.2 / 4.1 / 5.3	4
2. Welcher Konflikt?	1	Die Aufgabenstellung: “Ein bis zwei vollständige Sätze“ kann auf zwei Konflikte hindeuten und kann verwirrend sein. Warum die Aufgabe nur mit einem Punkt bewertet wird, ist nicht ganz nachvollziehbar	1.2 / 4.1 / 5.3	3
3. Pinien und Olivenbäume	2	Multiple-Choice-Wahl mit einer richtigen Aussage.	1.2 / 4.1 / 5.3	4

[5] Ratingskala: 1 = keine Übereinstimmung; 2 = geringe Übereinstimmung; 3 = weitgehende Übereinstimmung; 4 = völlige Übereinstimmung

4. Waldbrand	3	3 Gründe in Stichworten	1.2 / 4.1 / 5.3	4
5a. Geschichte als Einmischung in den Streit 5b. Reaktion der Leute	2	Antworten in ganzem Satz	1.2 / 4.1 / 5.3 1.2 / 4.3 / 5.3	3 3
Total Punktzahl	10			21/6=3.5
Wortschatz				
6. Synonyme	2	Zwei Synonyme der gleichen Wortart. Die Aufgabe fordert die Kenntnis des Begriffs „Synonym“ und der entsprechenden Wortart	1.1 / 1.3	3
7. Antonyme	2	Der Begriff „Antonym“ wird mit dem Klammerausdruck (Gegenteil) erklärt.	1.1	3
8. Bedeutung eines Wortes	1	Die Bedeutung eines Wortes muss in einem Satz erklärt werden.	1.1	4
9. Zwei Vergleiche	2	Als Hilfe wird ein Beispielsvergleich genannt.	1.1 / 1.2	4
10. Wörter der gleichen Wortfamilie	4	Die Aufgabe bedingt neben der Wortschatzkenntnis auch die Wortartenkenntnis	1.1 / 1.2/ 1.3	2
Total Punktzahl	11			16/5=3.2
Grammatik, Interpunktion, Orthografie		Keine Aufgabe zur Interpunktion und zur Orthografie		
11. Indirekte Rede	21/2	Kein Hinweis darauf, dass die indirekte Rede im Konjunktiv stehen muss.	1.3	2
12. Fallbestimmung	11/2	3 Fälle gesucht	1.3	4
13. Passiv	2	Kein Hinweis darauf, ob die Zeitform beibehalten werden muss.	1.3	3
14. Satzgliedbestimmung	3	Wahlmöglichkeit: Von vier Satzgliedern müssen drei gewählt werden.	1.3	4
15. Grundwortarten mit zusätzlichen Angaben	3	Wortarten mit genauen Angaben bestimmen. Angaben werden gemacht	1.3	4
Total Punktzahl	12			17/5=3.4
Text verfassen				
Ein Ereignis: „Mit dem Feuer spielen“	16	Im Auftrag wird erwähnt, dass „mit dem Feuer spielen“ im übertragenen Sinn gebraucht werden muss. Der Auftrag enthält, drei Fragen zur inhaltlichen Unterstützung und eine formale Vorgabe: ganze Sätze, 120 bis 150 Wörter.	3.1 / 3.2 / 3.3	3

		Keine Hinweise auf Anforderungen Beurteilungskriterien und deren Gewichtung, wie kommunikative Dimension, (Absicht; Adressat); inhaltliche Dimension, (Wie); formale Dimension (Darstellung); sprachliche Dimension (Stil, Richtigkeit)		
Total Punktzahl	16			3

Tabelle 4: Aufnahmeprüfung 2010 Deutsch

Aufgabe	Max. Punkt-zahl	Kommentar zur Aufgabe durch Prüfungsbeurteiler	Übereinstimmung Lehrplanziel(e)	Rating (1-4)
Textverständnis				
1. Wirkung des Kaufmanns	1	Die Antwort-Darstellung mit einer Zeile über der Frage und zwei Zeilen darunter ist verwirrlich. „Wie wirkte der Kaufmann?“ Die Frage ist offen formuliert. Es gibt keine Angabe, ob da die Wirkung am Anfang oder am Schluss gemeint ist. Wird in der Lösung: „nett und reisserisch“ akzeptiert?	1.2 / 4.1 / 5.3	2
2. Erklärung in eigenen Worten	2	Es gibt 7 Möglichkeiten, die er „kauft“. Wie viele müssen erklärt werden?	1.2 / 4.1 / 5.3	3
3. Krankheit	1	„Krankheit“ Anführungs- und Schlusszeichen weisen auf eine eigenartige Krankheit hin	1.2 / 4.1 / 5.3	3
4. Absicht der Geschäftspartner	1	Offene Frage, die aus dem Zusammenhang beantwortet werden muss.	1.2 / 4.1 / 5.3	3
5. Möglichkeiten des Betrügens	1	Es sind Verständnisfragen, deren Antworten nicht im Text wörtlich zu finden sind, sondern selbst überlegt werden müssen.	1.2 / 4.1 / 5.3	3
6. Multiple-Choice	1	Richtige Antwort markieren, erfordert genaues Lesen der Antwortmöglichkeiten.	1.2 / 4.1 / 5.3	3
7. Kaufmann merkte Veränderung nicht	1	Er merkte zwar, dass etwas los war, er erkannte aber nicht seine persönliche Veränderung.	1.2 / 4.1 / 5.3	2
8. Textsorte	1	Bedingt Kenntnis der drei Textsorten: Parabel, Satire, Fabel. Parabel und Satire stehen nicht explizit im Lehrplan	1.2 / 4.1 / 5.3 / 6.3	2
Total Punktzahl	9	Zahl stimmt nicht mit der maximalen Punktzahl		21/8

		auf dem Titelblatt überein.		2.6
Wortschatz, Stil				
9. Wortfamilie: Beruf und Tätigkeit	2	Der Begriff Wortfamilie muss bekannt sein. Beispiel hilft bei der Lösung	1.1	3
10. Gegensatzpaare	3	Die Lösung ist nicht bekannt, aber es ist schwierig, wirkliche Gegensatzpaare zu finden: sprechen – zuhören? Ruhig bleiben – erschrecken. Sind das wirkliche Gegensätze? Aufgabe ist zeitaufwändig, da 9 Zeilen dreimal gelesen werden müssen.	1.1	3
11. Wortfamilien	3	In der Aufgabenstellung wird nicht explizit darauf verwiesen, dass die Tabelle mit anderen Wortarten derselben Wortfamilie zu ergänzen ist.	1.1 / 1.3	2
12. Synonyme	2	Der Begriff Synonyme muss bekannt sein.	1.1 /1.3	3
13. Satzverbindung	2	Recht anspruchsvoll. Die Pluralform im Auftrag, Pronomen – Konjunktionen kann verwirrend sein, da ein Pronomen und eine Konjunktion gesucht wird	1.1 / 1.3	3
Total Punktzahl	12	Obwohl der Abschnitt „Wortschatz, Stil“ bezeichnet ist, ist einiges Grammatikwissen notwendig: Wortarten, Pronomen, Konjunktionen.		14/5 2.8
Grammatik, Interpunktion				
14. Steigerung zweier Adjektive	1		1.3	4
15. Bedeutung von Modalverben	2	Die Bedeutung von Modalverben ist bei Lernenden nicht explizit bekannt. „lassen“ ist kein Modalverb.	1.1 / 1.3	1
16. Passiv	1	Kein Hinweis darauf, ob die Zeitform beibehalten werden muss.	1.3	3
17. Direkte Rede	1	Hinweis auf Interpunktion	1.3	4
18. Kommasetzung	2	In diesen zwei Sätzen mit zwei Kommas stecken drei Kommaregeln.	3.3	4
19. Wortbestimmung	3	Kein Hinweis auf genaue Bestimmungen wie Wortart, Modus, Zahl, etc.	1.3	3
20. Fallbestimmung	2	Fälle von Satzgliedern bestimmen. Beim 2. Satzglied ist nur ein Teil des ganzen Satzgliedes „das Gesicht eines Tigers“ unterstrichen.	1.3	4
Total Punktzahl	12	Von den 12 Punkten sind zwei Punkte für die		23/7

		Interpunktion und 10 Punkte für Grammatik zu erreichen.		3.3
Textschaffen				
Ein Ereignis: „wahres Gesicht“ zeigen	16	Im Auftrag wird nicht explizit erwähnt, dass „wahres Gesicht“ im übertragenen Sinn gebraucht werden muss. Zwei Fragen zur inhaltlichen Unterstützung und eine formale Vorgabe: 120 bis 150 Wörtern. Keine Hinweise auf Anforderungen, Beurteilungskriterien und deren Gewichtung, wie kommunikative Dimension, (Absicht;Adressat) inhaltliche Dimension, (Wie) formale Dimension (Darstellung) und sprachliche Dimension (Stil, Richtigkeit)	3.1 / 3.2 / 3.3	3
Total Punktzahl	16			3

Aus der Übersicht wird ersichtlich, dass Grobziel 1.2 „Genaues Lesen und Verstehen“ und Grobziel 1.3 „Sprachliche Strukturen“ am häufigsten geprüft werden.
Die Grobziele 2.1 bis 2.5 „Mündliche Kommunikation“ kommen in den Prüfungen nicht vor.

Tabelle 5: Zielabdeckung mit dem Lehrplan

Übereinstimmung mit Grobzielen			ohne Übereinstimmung mit Grobzielen	
Grobziel	Anzahl	Grobziel	Grobziel	Grobziel
1.1	11x	Bedeutungsbreite von Wörtern	2.1	Situationsgerechte mündliche Kommunikation
1.2	16x	Erfahrungen und Empfindungen sprachlich genau erfassen	2.2	Auf die Partnerin / den Partner eingehen, Hörverständnis
1.3	16x	Denkstrukturen und sprachliche Strukturen	2.3	Mündliche Textformen, Diskussionen führen
3.1	2x	Angemessene schriftliche Kommunikation	2.4	Gepflegter Ausdruck in Mundart und Standardsprache
3.2	2x	Text ordnen und strukturieren	2.5	Besonderheiten der eigenen Mundarten
3.3	2x	Rechtschreibung und Zeichensetzung, grammatikalische Normen	4.2	Wort- und sprachbegleitende Kommunikation
4.1	13x	Dimension des Verstehens	4.3	Nonverbale Mitteilungsmöglichkeiten
5.3	14x	Verschiedene Lesetechniken	5.1	Die einzelnen Medien kennenlernen

6.1	2x	Auseinandersetzung mit sprachlichen Kunstwerken	5.2	Eigenheiten fiktionaler und nicht-fiktionaler Texte
6.3	2x	Techniken literarischen Schaffens	5.4	Subjektivität jeder Information / Manipulation
			5.5	Kritischer Umgang mit Medien
			6.2	Hintergründe und Werte sprachlicher Kunstwerke
			6.4	Eigenes sprachliches Gestalten
			7.1	Sprache in zwischenmenschlicher Beziehung
			7.2	Taktvoller Sprachgebrauch, Konfliktlösung
			7.3	Verantwortung gegenüber der Sprache

3.1.3. Prüfungsvergleich mit den Grundkompetenzen

Der nachfolgende Abschnitt berücksichtigt die Nationalen Bildungsstandards, die von der EDK-Plenarsversammlung am 16. Juni 2011 freigegeben wurden.

Das Kompetenzmodell „Schulsprache“ unterscheidet zwischen vier Kompetenzbereichen: Hören, Lesen, Sprechen und Schreiben. Im Hinblick auf die Konstruktion von Lehrplänen werden diese vier Kompetenzbereiche um zwei weitere ergänzt: Orthografie und Grammatik.

Die Aufnahmeprüfungen 2009 und 2010 setzen sich mit den Kompetenzen in Lesen, Schreiben, Orthografie und Grammatik auseinander.

Lesen

GRUNDKOMPETENZEN | LESEN | SCHULSPRACHE | 11. SCHULJAHR (ENDE SEKUNDARSTUFE I)
Die Schülerinnen und Schüler können vielfältige schriftliche Textarten verstehen, z.B. Kurzgeschichten und Romane (Jugendliteratur), Interviews, Reportagen, Anleitungen und argumentative Texte. Insbesondere können sie einen Text als Ganzes verstehen sowie zentrale inhaltliche Elemente erkennen und mit ihrer Lebenswelt in Verbindung bringen.
Sie können Textart und Textfunktion sowie die Autorintentionen erkennen. Sie können explizite und implizite Informationen erkennen und Schlussfolgerungen auf Basis einzelner Informationen oder ganzer Textteile ziehen sowie einen eigenen Standpunkt zum Text ausdrücken.
Sie sind fähig, ihr Textverständnis zu ergänzen, indem sie textstrukturierende Elemente (Inhaltsverzeichnis, Titel, Layout) und Illustrationen (Tabellen, Grafiken) nutzen.

Grundlage für das Textverständnis sind in beiden Prüfungen literarische Texte. Die Fragen zum Textverständnis zielen auf das Verstehen des Textes als Ganzes und auf das Erkennen zentraler inhaltlicher Elemente. Es geht um das Erkennen von Zusammenhängen auf Basis einzelner Informationen.
Damit decken sich die Fragen zum Textverständnis weitgehend mit den ersten beiden Grundkompetenzen zum Lesen.
Was fehlt, ist der Bezug zur eigenen Lebenswelt und das Ausdrücken eines eigenen Standpunktes und Beurteilen eines Textes.

Schreiben

GRUNDKOMPETENZEN | SCHREIBEN | SCHULSPRACHE | 11. SCHULJAHR (ENDE SEKUNDARSTUFE I)
Die Schülerinnen und Schüler können der jeweiligen Kommunikationssituation angemessene Texte verschiedener Genres schreiben (z.B. Erzählen, Argumentieren, Erklären, Informieren, Auffordern). Sie verwenden dazu die entsprechenden sprachlichen Muster und die in der Aufgabenstellung vorgegebenen Informationen zum Thema und zur Schreibsituation.
Sie können textstrukturierende Mittel (z.B. Titel und Absätze) verwenden und durch Textverknüpfungsmittel, insbesondere Konjunktionen, Zeigewörter (z.B. hier, dort) und Pronomen den Textzusammenhang deutlich machen (Kohäsion).
Sie können ihre Texte überarbeiten, indem sie Rechtschreibung, Grammatik und verwendete stilistische Mittel (z.B. passende Wortwahl und Register) überprüfen.

Die beiden Schreibaufträge sind Erzählungen, in der eine Redensart im übertragenen Sinne dargelegt werden muss, in die vorgegebene inhaltliche Informationen verwendet werden können.
Die Aufgabenstellung zum Texte schaffen deckt sich mit den Grundkompetenzen Schreiben.
Die Aufgabenstellung müsste allerdings differenzierter formuliert sein und auch Hinweise auf Funktion, Ziele, Adressaten und Adressatinnen machen, um die Kommunikationssituation des Schreibens zu klären.

Grammatik

GRUNDKOMPETENZEN | GRAMMATIK | SCHULSPRACHE DEUTSCH | 11. SCHULJAHR (ENDE SEKUNDARSTUFE I)
Die Schülerinnen und Schüler kennen die zentralen grammatischen Proben soweit, um damit einfach strukturierte Sätze zu analysieren. Zudem können sie diese für die Klärung orthografischer Probleme zumindest auf Aufforderung hin nutzen (analytische Grammatik- und Rechtschreibkompetenz).
In der Reflexion über Sprachstrukturen (z.B. in Situationen zu Sprachbewusstsein / Language Awareness / Eveil aux langues) können sie ihr Wissen über

Wortbedeutungen in Fremdsprachen zur Analyse von sprachlichen Phänomenen nutzen.
Sie können zudem schulgrammatische Begriffe zur Beschreibung des von ihnen erkannten sprachlichen Phänomens nutzen. Sie können die Reflexionen bzw. die Beschreibungen der anderen nachvollziehen.

Die Aufgaben zur Grammatik verlangen das losgelöste Bestimmen und Umsetzen von grammatischen Phänomenen. Sie dienen nicht der Analyse, dem Sprachvergleich oder der eigentlichen Anwendung im eigenen Schreiben.

Orthografie

GRUNDKOMPETENZEN | ORTHOGRAFIE | SCHULSPRACHE DEUTSCH | 11. SCHULJAHR (ENDE SEKUNDARSTUFE I)
Die Schülerinnen und Schüler können die orthografischen Regeln soweit anwenden, dass in einem Text die vorhandenen Fehler ein flüssiges Lesen nicht behindern.
Sie kennen die relevanten Rechtschreibregeln und können sie in isolierten Situationen anwenden.
Für die Klärung von orthografischen Problemen nutzen sie grammatische Proben auf Aufforderung hin.
Komplexere Proben, vor allem syntaktische Proben, werden unter Anleitung durchgeführt.
Die Schülerinnen und Schüler wenden einzelne Strategien zum Erwerb (z.B. Memorierungsstrategien) und in isolierten Aufgaben zur Lösung von Rechtschreibproblemen (z.B. die Lehrperson bzw. einen Mitschüler oder eine Mitschülerin fragen, das Wörterbuch benutzen) an.
Sie können Korrekturstrategien auf Aufforderung hin und unter Anleitung anwenden, d.h. die Schülerinnen und Schüler kennen eine Anzahl Nachschlagewerke und können sie mehrheitlich erfolgreich nutzen.
Sie können in einem Forschungsauftrag orthografische Regeln ableiten.

Eine eigentliche Orthografie-Aufgabe kommt nur in der Prüfung 2010 vor, in der es gilt die Kommas richtig zu setzen. Anderseits spielt die Orthografie im Textschreiben eine wichtige Rolle, zu der Punkt 1 und 6 zählen.
Die Grundkompetenzen zur Rechtschreibung werden hauptsächlich in der Anwendung, im Bereich Text verfassen, geprüft.

3.1.4. Prüfungsvergleich mit dem Stoffplan

Im Stoffplan für die Aufnahmeprüfung wird auf das obligatorische Lehrmittel der Sekundarstufe I „Sprachwelt Deutsch“ verwiesen. Dieses Buch bildet die Grundlage für den Deutschunterricht auf der Sekundarstufe I und ist auf dem Lehrplan aufgebaut.

Die drei Bereiche: Texte schaffen, Lesen und Verstehen, Sprachsystem sind drei der vier Bereiche des Deutschunterrichtes. Es fehlt der Bereich: Hören und Sprechen.

Der Stoffplan für die Aufnahmeprüfung orientiert sich am Lehrplan und am obligatorischen Lehrmittel der Sekundarstufe I. Der Stoffplan ist für den Leser, Lehrperson und Schüler / Schülerin, verständlich und nachvollziehbar verfasst.

Die Prüfungsaufgaben sind Bestandteile der im Stoffplan beschriebenen Bereiche.

Im Stoffplan steht unter „Texte schaffen“: Kriterien für situations- und adressatengerechtes Schreiben: Kohärenter Aufbau, relevanter Inhalt, sprachliche Korrektheit, angemessener Stil.
Die Aufgabenstellung zum Textschaffen in den Prüfungen 2009 und 2010 geben keine Hinweise auf die Schreibsituation und Adressaten.

3.1.5. Schlusskommentar

Prüfung allgemein
Die Aufnahmeprüfungen Deutsch 2009 und 2010 umfassen die Bereiche: Textverständnis; Wortschatz; Grammatik, Interpunktion, Orthografie; Text verfassen. Sie sind als schriftliche Prüfungen gestaltet.
Damit wird ein breites Spektrum des Deutschunterrichtes abgedeckt, aber es fehlt der Bereich „Mündlichkeit“: Hören und Sprechen.
Für die gesamte Prüfung stehen 80 Minuten zur Verfügung, wobei keine Zeitangaben für die einzelnen Bereiche angegeben werden. Damit ist die Zeit eher knapp bemessen, gilt es doch zum Textverständnis den literarischen Text als Ganzes zu lesen und zur Beantwortung der Fragen, Zeilen zu suchen und Inhalte mehrmals zu lesen, für Wortschatz und Grammatik sind weitere 10 (Prüfung 1009) resp. 13 (Prüfung 2010) Fragen zu beantworten, deren Kontext über Zeilenhinweise im Text zu suchen sind.
Die Prüfungsanlage in den Bereichen Wortschatz und Grammatik mit 23 resp. 24 Wertungspunkten ist auf Fragen aufgebaut, die mit richtig oder falsch gewertet werden können. Sprache wird dabei auf einen normativen Aspekt reduziert und hat nur wenig mit dem eigentlichen Sprachhandeln zu tun.
Kompetenzen der Sprache sind aber für das Sprachhandeln aufgebaut, auf das Situieren, Planen, Realisieren, Evaluieren und Reparieren von Textproduktion und Textrezeption. Dies wird in der Prüfungsanlage im Besonderen im „Text verfassen“ und zum Teil im „Textverständnis“ geprüft (27/26 Wertungspunkte).
Die Bewertung von 24 Punkten für Wortschatz / Grammatik und 26 Punkten für Textverständnis und Textproduktion ist mehr oder weniger gleichwertig. Wenn

man bedenkt, dass Rechtschreibung und Grammatik auch im Textschaffen bewertet wird, so erhält der normative Aspekt eine Überbewertung.

Textverständnis
Das Textverständnis basiert auf zwei literarischen Texten, wie sie in der Sekundarstufe I für das Niveau A und B üblich sind. Die Fragen erfassen ein Textverständnis, das auf textbezogenes Interpretieren hinzielt und somit einer wichtigen Kompetenz der PISA-Lesekompetenz entspricht.

Wortschatz und Grammatik
Die Aufgaben in Wortschatz und Grammatik beziehen sich auf die literarischen Texte, wozu auch Kontextverstehen hilfreich sein kann. Die Aufgaben sind nicht auf spezifische formale oder inhaltliche Merkmale der Texte ausgerichtet, sondern fragen nach textunabhängigem Wortschatz- und Grammatikwissen.
In der Aufnahmeprüfung von 2010 beschreiben einzelne Aufgaben nur ungenau die Anforderungen (Aufgaben 11 und 19) und erfordern, dass Schüler und Schülerinnen schon ähnliche Aufgaben gelöst haben, um die genaue Aufgabenstellung zu verstehen.

Texte schaffen
Die Aufgabenstellung im Textschaffen ist teilweise undifferenziert, sie gibt keine Hinweise auf die genaue Funktion, auf Ziele und Adressaten des Schreibens und gibt keine Angaben auf die Gewichtung der einzelnen Textaspekte (Inhalt, Stil, Orthografie)
Ein Text „aus dem Leben gegriffen“, wie in der Prüfung 2010, oder ein Text zu „einem eigenen Ereignis“, wie in Prüfung 2009, sind unterschiedlich schwierig. Ein Text „aus dem Leben gegriffen“ ist sicher schwieriger zu taxieren. In der Punktzahl werden diese beiden Textaufgaben unterschiedlich bewertet, das „Textschaffen“ in der Prüfung 2009, eigenes Ereignis, erhält 17 Punkte, das Textschaffen der Prüfung 2010, „ein Text aus dem Leben gegriffen“ erhält 16 Punkte. Damit wird der Schreiber mit einem „eigenen Ereignis“ bevorzugt.
Als formale Vorgabe wird im Textschaffen die Anzahl Wörter vorgegeben, 120 bis 150. Die Anzahl der Wörter ist dabei zu zählen und am Schluss hinzuschreiben, was sehr zeitaufwändig und allenfalls auch hinderlich für den Schreibfluss ist.

3.1.6. Fazit

- Die inhaltliche Validität der Prüfungsaufgaben stimmt weitgehend mit den Lernzielen des Lehrplanes Deutsch überein.
- Die Prüfung deckt den Sprachlernbereich der Mündlichkeit: Hören und Sprechen, nicht ab.

- Sprachliche Normen und sprachliche Strukturen sind im Verhältnis zum eigentlichen Sprachhandeln stärker gewichtet.
- Die Gewichtung der einzelnen Sprachbereiche ist ausgeglichen: Wortschatz und Grammatik werden mehr oder weniger gleich stark wie das Textverständnis und das Textschaffen bewertet. Diese Gewichtung deckt aber nicht die Grundkompetenzen ab, in denen das Sprachhandeln im Vordergrund steht.
- Die Aufgabenstellungen bedingen eine Vorbereitung mit ähnlichen Prüfungen, um Anforderungen und Gewichtung zu verstehen.
- Die Prüfungszeit ist eher knapp bemessen und lässt dem Schreiben als Prozess mit Ideen finden, entwerfen, überarbeiten, zu wenig Zeit.
- In den beiden Prüfungen 2009 und 2010 werden die einzelnen Bereiche (Textverständnis, Wortschatz, Grammatik/Interpunktion, Textschaffen) unterschiedlich bewertet (Textschaffen 2009: 17 Punkte; Textschaffen 2010: 16 Punkte).
- Die Aufgabenformulierungen und damit verbundenen indirekten Hilfen sind in den beiden Prüfungen nicht identisch.[6]

3.2. Beurteilung der Aufnahmeprüfungen Mathematik von 2009 und 2010 (Reinhard Hölzl und Marco Hübner)

3.2.1. Gewichtung (Punktevergabe)

Die Gewichtung der Aufgaben folgt einem strikten homogenen Ansatz: Jede Aufgabe ist mit gleich vielen Punkten (2) dotiert, unabhängig vom jeweiligen Schwierigkeitsgrad und Anspruchsgehalt. Die Korrekturlösungen der Aufgaben (zugänglich gemacht von Marco Hübner) haben „monolithischen" Charakter, d. h. es wird (implizit) nicht nur die Lösung der Aufgabe bepunktet, sondern auch das vom Aufgabensteller intendierte Verfahren. Systematisches Probieren anhand einer eigens erstellten Tabelle ist beispielsweise in A6 aus Arithmetik/Algebra 2009 kein bewertungsrelevantes Verfahren, da (implizit) eine Gleichung erwartet wird. Mit der zeitgemässen didaktischen Forderung nach multiplen Lösungswegen verträgt sich dieses Format nicht besonders.

3.2.2. Curriculare Validität

Die Prüfungsaufgaben sind konform zu den Lernzielen des aktuellen S1-Lehrplans, sie halten sich an den Stoffplan für die Aufnahmeprüfungen. Zu bemerken ist lediglich, dass manche der Aufgaben, was den Schwierigkeitsgrad betrifft, als grenzwertig anzusehen sind. D. h. sie erfordern im algebraischen oder

[6] Aufgabe 15, 2009: „Bestimmen Sie die Grundwortarten der unterstrichenen Wörter im folgenden Satz. Geben Sie, wo möglich dazu noch genaue Angaben zu Zeit, Person, Untergruppe und / oder Modus."
Aufgabe 19, 2010: „Bestimmen Sie die unterstrichenen Wörter im folgenden Satz so genau wie möglich."

geometrisch-konstruktiven Bereich ein relativ hohes technisches Vermögen (Beispiel A5 in Arithmetik/Algebra 2009 bzw. A2 in Geometrie 2010).

3.2.3. Kompetenzanalytische Beurteilung

Die Frage der kompetenzanalytischen Ausrichtung der Prüfungen im Sinne des Referenzrahmens von HarmoS hat uns am meisten interessiert. Wir haben jede Aufgabe nach den erforderlichen Kompetenzen beurteilt. In den nachfolgenden Tabellen ist das Ergebnis dieser Überprüfung dargestellt. Als Titel erscheint immer der hauptsächlich beanspruchte Kompetenzaspekt, ebenfalls dazugehörige „Nebenaspekte" sind in der Beschreibung erwähnt. Um ein Kompetenzprofil zu erstellen, haben wir die Anzahl Nennungen von Haupt- und Nebenaspekten in einem Balkendiagramm dargestellt – gewissermassen als Kompetenzprofil der Prüfung.
Hinweis: Den Kompetenzaspekt Verwenden von Instrumenten und Werkzeugen haben wir nicht aufgeführt, da er in der Geometrie (Zirkel, Lineal) bzw. Arithmetik (Taschenrechner) als selbstverständlich erscheint.

Tabelle 6: Aufnahmeprüfung 2009 – Geometrie

Aufgabe	Kompetenzaspekt
A1	Mathematisieren & Modellieren Eine Alltagssituation ist durch ein geometrisches Modell zu beschreiben (zentrale Grössen: Raumhöhe und Schrankdiagonale). Da die Kenntnis des Satzes von Pythagoras vorausgesetzt wird auch Wissen, Erkennen & Beschreiben. Problemlöseaspekt vorhanden, da die Pythagoraskonfiguration „hineingesehen" werden muss.
A2	Operieren & Berechnen Auf der Basis geometrischer Grundkonstruktionen ist eine anspruchsvollere Viereckskonstruktion zu leisten. Je nach Vorkenntnissen deutlicher Problemlösecharakter. Explizite Aufforderung zu Beschreibung des Konstruktionsweges: Darstellen & Kommunizieren.
A3	Wissen, Erkennen & Beschreiben Begriffs- und Satzwissen notwendig (Umfang, Fläche, Inhaltsformeln für Kreis und Dreieck). Operieren & Berechnen. Da die gegebene Konfiguration unterschiedliche gedeutet werden kann (z. B. weisse Fläche = Halbkreis), wird auch der Kompetenzaspekt Interpretieren (& Reflektieren der Resultate) angesprochen.
A4	Wissen, Erkennen & Beschreiben Begriffs- und Satzwissen notwendig (Quader, Prisma, Rauminhalte derselben). Operieren & Berechnen. Da aus der Zeichnung relevante geometrische Informationen entnommen werden müssen, wird auch der Kompetenzaspekt Darstellen (& Kommunizieren) angesprochen.
A5	Operieren & Berechnen / Wissen, Erkennen & Beschreiben Winkelberechnungen auf der Grundlage elementarer Sätze der ebenen Geometrie (Satz des Thales, Basiswinkel in gleichschenkligen Dreiecken) und vorhandener Kenntnisse über Innenwinkelsummen bei Drei- und Viereck. Da diese Begriffs- und Satzkenntnisse nicht einfach abgefragt werden, sondern in der gegebenen Konfiguration erkannt werden müssen, ist damit auch ein Problemlöseaspekt vorhanden.

Tabelle 7: Aufnahmeprüfung 2010 – Geometrie

Aufgabe	Kompetenzaspekt
A1	Operieren & Berechnen Mehrschrittige Berechnungsaufgabe, in der aus einer Flächenbeziehung auf eine Streckenlänge zu schliessen ist. Da die Kenntnis des Satzes von Pythagoras vorausgesetzt wird auch Wissen, Erkennen & Beschreiben. Problemlöseaspekt etwas vorhanden.
A2	Operieren & Berechnen Auf der Basis geometrischer Grundkonstruktionen ist eine anspruchsvollere Konstruktion zu leisten. Fehlende Angaben sind zunächst zu berechnen. Begriffs- und Satzwissen notwendig: Symmetrieeigenschaft der Winkelhalbierenden, Thaleskreis als geometrischer Ort: Wissen, Erkennen & Beschreiben. Explizite Aufforderung zu Beschreibung des Konstruktionsweges: Darstellen & Kommunizieren. Erhöhte Anforderungen an das Textverständnis.
A3	Operieren & Berechnen Umfang und Inhalt eines Kreisrings sind zu bestimmen. Da die Konfiguration nicht als Zeichnung vorliegt, sondern aus einer verbalen Beschreibung vorgängig erstellt werden muss, ist auch der Kompetenzaspekt Darstellen (& Kommunizieren) angesprochen. Erhöhte Anforderungen an das Textverständnis.
A4	Operieren & Berechnen Aus der Oberfläche eines Quaders ist auf dessen Rauminhalt zu schliessen. Identifikation gesuchter und gegebener Grössen in einer vorgegebenen Skizze. Kompetenzaspekte Wissen, Erkennen & Beschreiben sowie Darstellen (& Kommunizieren) geringgradig ausgeprägt.
A5	Operieren & Berechnen Das Vorstellungsvermögen relativ stark fordernde kopfgeometrische Problemstellung. Perspektivische Zeichnung und Körperabwicklung sind in Einklang zu bringen. Eine Überprüfung ist zwar nicht explizit gefordert, wird aber durch die einzuzeichnenden Strecken nahegelegt, daher auch (implizit) Interpretieren & Reflektieren der Resultate.

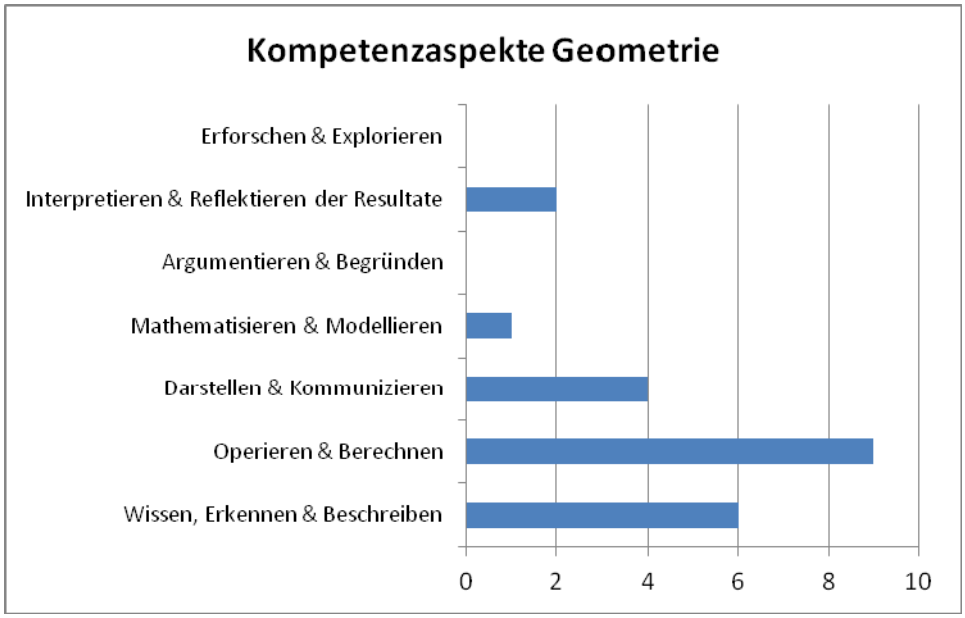

Abbildung 1: Anzahl Nennungen des jeweiligen Kompetenzaspekts in den Prüfungen 2009 und 2010

Fazit: Die Aufnahmeprüfungen 2009/10 – Geometrie stimmen in ihrem Aufbau und Aufgabencharakter grundsätzlich überein, wobei die Prüfung 2010 sowohl in Bezug auf die angesprochenen Kompetenzaspekte als auch den Problemgehalt der Aufgaben und ihrer Konstruktion ausgewogener erscheint. Der am meisten beanspruchte Kompetenzaspekt stellt das Operieren & Berechnen dar, wobei hier zum Teil auch ein erhöhter Problemlöseaspekt zu beobachten ist. Grundsätzlich ist der Aufgabencharakter wie in den Aufnahmeprüfungen Arithmetik und Algebra geschlossen und weist ein eher geringes Differenzierungspotential im Hinblick auf multiple Lösungswege auf. Die curriculare Passung ist gegeben, allerdings erachten wir manche der Konstruktionsaufgaben als grenzwertig für die Sekundarstufe I.

Tabelle 8: Aufnahmeprüfung 2009 – Arithmetik/Algebra

Aufgabe	Kompetenzaspekt
A1	Operieren & Berechnen Eine pseudo-quadratische Gleichung ist zu lösen. Erforderliche Termumformungen auf der Basis des Distributivgesetzes und binomischer Formel.
A2	Operieren & Berechnen Fehlende Exponenten sind zu ergänzen, da diese zunächst probeweise eingesetzt werden können und ein Abgleich mit der linken und rechten Seite erforderlich ist, wird in der Aufgabe auch der Kompetenzaspekt Erforschen & Explorieren angesprochen.
A3	Mathematisieren (& Modellieren) Zu bilden sind Terme für ein figuriertes Zahlenmuster. Die zugrunde liegende Übersetzungsleistung spricht nur den Aspekt Mathematisieren an. Eine Modellierungsleistung ist hier nicht erforderlich.
A4	Operieren & Berechnen Stark umwandlungstechnisch orientierte Behandlung von Termen aus Summen und Produkten (Quotienten).
A5	Operieren & Berechnen Vereinfachung eines Bruchterms, eher grenzwertig für die Sekundarstufe I. Konzentration auf Technik: Ausklammern erkennen, Kürzen. Teilaufgabe b) ist auf den Einsetzungsaspekt abgestimmt (Variable als Platzhalter).
A6	Mathematisieren (& Modellieren) Anwenden der Zinsformel, in Teilaufgabe a) ist der Zins unbekannt, in b) die Zeit. Das mathematische Modell liegt also schon vor, die Übersetzung ist formelmässiger Natur, daher ist der Aspekt „Mathematisieren“ in einiger Nähe auch zu Operieren & Berechnen zu sehen.
A7	Mathematisieren (& Modellieren) Der vorhandene (Pseudo-)Kontext ist mit mathematischen Mitteln zu beschreiben, die Gleichung als mathematisches Modell der Sachsituation ist zu lösen, daher auch Operieren & Berechnen. (Insbesondere auch vor dem Hintergrund, dass in den Korrekturlösungen keine Ersatzstrategien wie z. B. systematisches Probieren, Tabellen, zugelassen sind.)
A8	Mathematisieren (& Modellieren) Zweischrittige Weg-Geschwindigkeit-Aufgabe, dadurch auch Bezug zum Inhaltsbereich Grössen & Masse. Die Mathematisierung ist wie in A6 formelmässiger Natur. Da eine Skizze den Lösungsprozess

	unterstützt (allerdings nicht gefordert wird) ist zu einem kleinen Teil auch der Kompetenzaspekt Darstellen (& Kommunizieren) angesprochen.

Tabelle 9: Aufnahmeprüfung 2010 – Arithmetik/Algebra

Aufgabe	Kompetenzaspekt
A1	Operieren & Berechnen Die Gleichwertigkeit vorgegebener Terme ist in Bezug auf zwei Resultate zu beurteilen. Da es sich um kein sogenanntes „strukturiertes Päckchen“ handelt, ist zur Lösung jeder der vorgegebenen Terme zu vereinfachen.
A2	Operieren & Berechnen Eine pseudo-quadratische Gleichung ist zu lösen. Erforderliche Termumformungen auf der Basis des Distributivgesetzes. Bruchterme mit ganzen Zahlen im Nenner.
A3	Darstellen (& Kommunizieren) Ein Summenprodukt ist einer aufgeteilten Rechtecksfläche zuzuordnen. Anschliessend ist der als Teilrechteck zu interpretierende Inhalt als Summe weiterer Teilrechtecke darzustellen. Darstellen von Termen als Flächeninhalte.
A4	Operieren & Berechnen Stark umwandlungstechnisch orientierte Behandlung von Termen aus Summen und Produkten (Quotienten).
A5	Operieren & Berechnen Vereinfachung eines Bruchterms, eher grenzwertig für die Sekundarstufe I. Konzentration auf Technik: Potenzgesetze und Bruchdivision.
A6	Mathematisieren (& Modellieren) Anwenden der Zinsformel in einem zusammengesetzten Sachzusammenhang aus Teilzinssätzen und Gesamtzinssatz. Da eine Gleichung zu lösen ist auch Operieren & Berechnen.
A7	Mathematisieren (& Modellieren) Der vorhandene (Pseudo-)Kontext ist mit mathematischen Mitteln zu beschreiben, die Gleichung als mathematisches Modell der Sachsituation ist zu lösen, daher auch Operieren & Berechnen. (Vgl. Anmerkung zur A7 in der Prüfung 2009.)
A8	Mathematisieren (& Modellieren) Mehrschrittige Weg-Geschwindigkeit-Aufgabe, dadurch auch Bezug zum Inhaltsbereich Grössen & Masse. Die Mathematisierung ist formelmässiger Natur. Da eine Skizze den Lösungsprozess unterstützen würde (allerdings nicht gefordert wird) ist zu einem kleinen Teil auch der Kompetenzaspekt Darstellen (& Kommunizieren) angesprochen. (Aus der Korrekturlösung ist ersichtlich, dass auch ein Antwortsatz verlangt wird.)

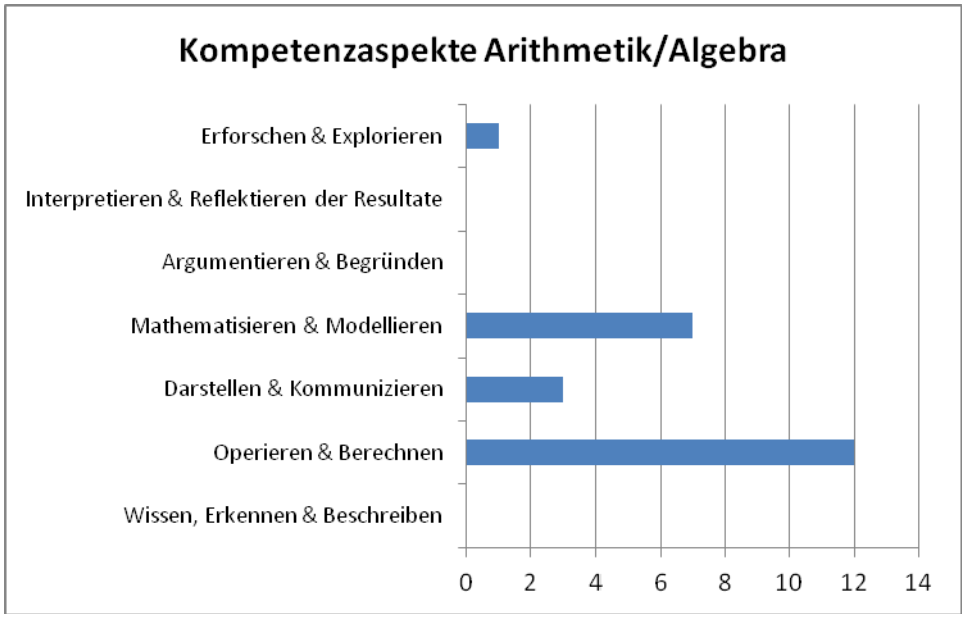

Abbildung 2: Anzahl Nennungen des jeweiligen Kompetenzaspekts in den Prüfungen 2009 und 2010

Fazit: Die Aufnahmeprüfungen 2009/10 – Arithmetik/Algebra stimmen in ihrem Aufbau und Aufgabencharakter überein. Sie sprechen nahezu ausnahmslos die Kompetenzaspekte Operieren & Berechnen sowie Mathematisieren an, wobei zu Letzterem meist das mathematische Modell als zu verwendende Formel gegeben ist. Die Aufgaben konzentrieren sich auf die technische, verfahrensorientierte Seite mathematischer Performanzdarstellung. Es handelt sich um geschlossene Aufgaben mit geringem Differenzierungsgehalt und teils tradierten Pseudokontexten.

3.3. Beurteilung der Aufnahmeprüfungen Englisch von 2009 und 2010 (Monika Mettler und Katharina Fischer)

3.3.1. Übereinstimmung mit dem Lehrplan

Die Analyseergebnisse zur Übereinstimmung mit dem Lehrplan werden im Folgenden in Tabellenform dargestellt.

Tabelle 10: Beurteilung der Übereinstimmung mit dem Lehrplan Englisch

Kriterium	Rating[7]				Bemerkungen
Schwierigkeitsgrad	1	2	3	4 x	Der aktuelle Zentralschweizer Lehrplan Englisch von 2004 setzt vier Jahre Englischunterricht auf der Primarstufe voraus. Im Kanton Luzern werden erstmals im Schuljahr 2013/14 Schüler/innen der 3. Sekundarstufe an die Aufnahmeprüfungen vom Frühjahr 2014 kommen, die nach diesem Lehrplan unterrichtet wurden. Die Analyse der Aufnahmeprüfungen 2009 und 2010 basiert daher noch auf dem Lehrplan 1996. Bemerkung: Schüler/innen der Kantone NW, OW; UR, SZ und ZG hatten in ihrer Primarschulzeit schon Englischunterricht und werden an den Aufnahmeprüfungen von 2012 und 2013 entsprechend andere Lernvoraussetzungen haben. Für die Prüfungen im Frühjahr 2014, werden die Voraussetzungen dann für alle Schüler/innen gleich sein.
Mehrteiligkeit			x		Die Berücksichtigung verschiedener kommunikativer Fertigkeiten (skills) in einer Englischprüfung ist ein fachdidaktisches Gütekriterium. In den Aufnahmeprüfungen 2009 und 2010 werden drei Kompetenzen Hörverständnis, Leseverstehen und Textproduktion geprüft. Jedoch wird die mündliche Produktion (Speaking) nicht geprüft. Obwohl organisatorisch aufwändig, wäre es wünschenswert, wenn auch diese Kompetenz geprüft würde.
Kontextualisierung von			x		Es fällt positiv auf, dass die Aufnahmeprüfungen Grammatik

[7] Ratingskala: 1 = keine Übereinstimmung; 2 = geringe Übereinstimmung; 3 = weitgehende Übereinstimmung; 4 = völlige Übereinstimmung

Grammatik und Wortschatz					und Wortschatz im Kontext der kommunikativen Fertigkeiten prüfen. Es wird jedoch nicht ersichtlich, wie das Writing korrigiert wird und welches Gewicht dabei grammatische Korrektheit und Wortschatzumfang erhalten.
Gewichtung der Teile			x		In der Prüfung 2010 sind die drei Teile gleich gewichtet, in derjenigen von 2009 erhält das Leseverstehen am meisten Gewicht (24 von 61 Punkten) und Hörverstehen am wenigsten (17 von 61 Punkten). Eine gleichwertige Gewichtung wird sehr begrüsst.
Lebensweltbezug und Realitätsnähe				x	Die Alltagsrelevanz von Aufgaben hat einen direkten Zusammenhang mit der Motivation. Alltagsrelevanz zeigt sich auch an einer möglichst hohen Authentizität der Aufgabe.
Testobjektivität				x	Testanweisungen sind jeweils in EN, damit wird immer auch Leseverstehen mitgeprüft. Jedoch sind die Anweisungen einfach und klar gehalten. Es kann daher vorausgesetzt werden, dass die Sch die Anweisungen verstehen. Die Testformate dürften den Sch von Prüfungsvorbereitungen her bekannt sein. Es handelt sich um Standardformate, wie sie in den Cambridge Prüfungen verwendet werden.
Bezug zu Standards Harmos				x	Die Standards für Harmos Fremdsprachen wurden noch unter dem „alten System“ (d.h. Beginn des Englischunterrichts ab 7. Schuljahr) entwickelt und können daher für die vorliegende Analyse beigezogen werden. Gemäss Harmos sind am Ende des 9. Schuljahres die Grundkompetenzen wie folgt festgelegt: Hören/Lesen: A2.2 Sprechen: A2.2 Schreiben A2.1. Gute Schüler/innen sollten entsprechend also auf einem etwas höheren Niveau sein. Die Aufgaben der vorliegenden Aufnahmeprüfungen werden wie folgt eingestuft: Reading: B1 Listening: A2.2 bis B1, besonders jeweils der 2. Teil des Listening. Schreiben: A2.2., abhängig von den Bewertungskriterien (siehe Kriterium 4)

3.3.2. Globaler Eindruck

Die Gesamtbeurteilung der Aufnahmeprüfungen fällt positiv aus. Der Schwierigkeitsgrad scheint für Sek A-Schüler/innen angemessen. Die Themen sind relevant, lebensnah und interessant. Die Schreibaufgabe bietet Wahlmöglichkeiten. Die erlaubten Hilfsmittel (zweisprachiges Wörterbuch) sind richtig.

3.4. Beurteilung der Aufnahmeprüfungen Französisch von 2009 und 2010 (Hans-Peter Hodel)

Diese Beurteilung hat zum Ziel, die Übereinstimmung der vorhandenen Aufnahmeprüfungen BMS / FMS / WMS Französisch von 2009 und 2010 (in der Folge als „Prüfungen" bezeichnet) mit den Lernzielen des aktuellen Lehrplans, insbesondere auch mit den Niveaus gemäss Gemeinsamen europäischen Referenzrahmen (GER) zu prüfen, sowie weitere Qualitätsmerkmale validen Testens von Sprachkompetenzen (wie beispielsweise Unabhängigkeit der Prüfungsteile, Transparenz der Kriterien) zu untersuchen.
Es folgen zuerst eine Auflistung relevanter Qualitätsmerkmale von Sprachprüfungen und eine Darstellung der Situation des Fachs Französisch auf der Ebene des Lehrplans, dann ein Hinweis auf die für diese Evaluation benutzten Prüfmethoden. Der Hauptteil widmet sich dann der Untersuchung der Prüfungen, wobei zuerst die Übereinstimmung mit Lehrplan und weitere Qualitätsmerkmale bezüglich jeder der Teilfertigkeiten (Hörverstehen, Leseverstehen, Textproduktion) in beiden Prüfungen (2009 und 2010), dann weitere allgemeine Merkmale der beiden Prüfungen analysiert werden[8]. Abschliessend folgt die Zusammenfassung der wichtigsten Befunde der Evaluation.

3.4.1. Qualitätsmerkmale validen Prüfens

Für die Evaluation der Französischprüfung sind folgende Qualitätsmerkmale validen Prüfens von sprachlich-kommunikativen Fertigkeiten im Fokus:

- die Übereinstimmung der Prüfungsinhalte mit dem Lehrplan (als ein Aspekt der Content Validitiy) – Prüfungsinhalte sollen den Lehrplan repräsentativ abbilden
- die Validitätsbereiche[9] des Konstrukts (Prüfung prüft effektiv die gewollte Kompetenz, Übereinstimmung Etikette-Inhalt) und Inhalt als Übereinstimmung der Prüfungsinhalte und –formate mit der Lernbiographie der KandidatInnen (=Prüfungsteilnehmenden)
- die Unabhängigkeit der Prüfungsteile voneinander

[8] Dabei werden positive wie negative Befunde beachtet, letztere wohl aber stärker fokussiert.
[9] consequential / predictive validity (Prognosequalität) sind hier nicht Gegenstand, wären aber auch interessant.

- eine ausreichende (≥30) Anzahl einzelne unabhängige Beobachtungen / Items / Punkte in der Rezeption[10] sowie eine ausreichende (= mind. zwei) Anzahl unterschiedlicher Texttypen / Themen in der Interaktion / Produktion[11]
- komplette und ausführliche Information über alle Belange der Prüfung zuhanden der KandidatInnen
- Klarheit, Logik, Konsistenz, Zugänglichkeit und Neutralität der Aufträge
- Aufträge, die realitätsnahe Kommunikation auslösen
- Festlegung der Bedingungen / Vorgaben der Durchführung der Prüfungen
- Definition der Evaluationskriterien, Lösungen, Punkteverteilung, Bestehensnormen, Berechnungsverfahren, Umrechnung auf Noten

3.4.2. Lernziele /Niveaus Ende Sekstufe I gemäss Lehrplan und HarmoS

Die Übereinstimmung mit dem Lehrplan ist ein zentrales Validitätsmerkmal. Angesichts der Bewegung, welche das Fach Französisch, vor allem durch die Einführung des Englisch in den letzten Jahren erfahren hat, soll dieses Merkmal ausführlicher dargestellt werden.
Der geltende IEDK Lehrplan Französisch aus dem Jahr 2000 bezieht sich auf das Fach Französisch als erste Fremdsprache (Beginn des Fachs ab 5. Primarklasse). Die Ziele für die Sekundarstufe I werden mit den Niveaus A2 (Grundanforderungen) und B1 (erweiterte Anforderungen) umschrieben. Die Niveaus stammen aus dem GER (Gemeinsamer europäischer Referenzrahmen für Sprachen). Bei den Grundanforderungen soll gemäss diesem Lehrplan das Gewicht besonders auf den Teilfertigkeiten Hören und Sprechen liegen, bei den erweiterten Anforderungen sei darauf zu achten, dass mit guten und sehr guten Schülerinnen und Schülern an möglichst allen Teilzielen (=Teilfertigkeiten) gearbeitet werde.
Inzwischen ist Französisch zweite Fremdsprache geworden.[12] Die im Lehrplan 2000 beschriebenen Ziele gelten jedoch nach wie vor, wie das Beispiel aus OW zeigt (Amt für Volks- und Mittelschulen OW Mai 2007[13]).

[10] siehe internationale Diplome wie TCF/CIEP ou TOEFL
[11] siehe GER Kap. 9.2.1
[12] Aus dem Kanton Luzern werden erstmals im Schuljahr 2013/14 Schüler/innen der 3. Sekundarstufe an die Aufnahmeprüfungen vom Frühjahr 2014 kommen, die Französisch als zweite Fremdsprache haben, aus NW und OW erstmals im Jahre 2012, aus UR, SZ und ZG im Jahre 2013.
[13] http://www.ow.ch/dl.php/de/46713bf824aa6/Stoffverteilung_envol_7-8_OS_OW.pdf (20120106)

Die EDK-Plenarversammlung zu den Nationalen Bildungsstandards (HarmoS) vom 16. Juni 2011[14] beschreibt die Grundkompetenzen für die erste und zweite Fremdsprache mit A2.2 (mit der Ausnahme von A2.1 für Schreiben –

Sprachliche Ziele gemäss der Schweizer Version des europäischen Sprachenportfolios ESP

Fertigkeiten / Schuljahr	Rezeptiv				Produktiv			
	Hören		Lesen		Sprechen		Schreiben	
	GA*	EA**	GA	EA	GA	EA	GA	EA
7	A 2.2	B 1.1	A.2.2	B 1.1	A 2.1	A 2.2	A 2.1	A 2.2
8	A 2.2	B 1.2	A.2.2	B 1.2	A 2.1	B 1.1	A 2.1	B 1.1
9	A 2.2	B 1.2	A.2.2	B 1.2	A 2.1	B 1.1	A 2.1	B 1.1

*GA: Grundanforderungen (Niveau B)
**EA: Erweiterte Anforderungen (Niveau A)

Abbildung 3: Sprachliche Ziele gemäss Schweizer Version des ESP

Deskriptoren zu diesen Niveaus befinden sich im Anhang). Dies entspricht gut dem geltenden IEDK Lehrplan Französisch 2000 (Grundanforderungen).

Den Grundkompetenzen entsprechendes Kompetenzprofil für das 11. Schuljahr

1. und 2. Fremdsprache	Referenzniveaus GER	A1.1	A1.2	A2.1	A2.2	B1.1	B1.2	B2
Kompetenzbereiche	Hörverstehen							
	Leseverstehen							
	Teilnahme an Gesprächen							
	Zusammenhängendes Sprechen							
	Schreiben							

Abbildung 4: Den Grundkompetenzen entsprechendes Kompetenzprofil für das 11. SJ

Somit kann davon ausgegangen werden, dass Prüfungen welche im März des 3. (d. h. letzten) Schuljahres der Sekundarstufe I stattfinden und welche eine Bandbreite von Niveau A2.2 (Grundanforderungen) bis Niveau B1.1 abdecken, inhaltlich, d. h. sowohl in Bezug auf den Lehrplan wie auch in Bezug auf den Lernstand der KandidatInnen, als valide bezeichnet werden können.

3.4.3. Angewandte Prüfmethoden

Zur Einschätzung der Niveaus der Prüfungen wurde die Methode des Vergleichs der vorhandenen Texte / Aufgaben mit den Deskriptoren des GER benutzt. Diese Methode ist üblich, müsste jedoch z. B. durch Einbezug weiterer Personen und der Prüfungsergebnisse ergänzt werden[15]. Für weitere Qualitätsmerkmale wurde die Methode der konkreten Durchführung der Prüfungsteile durch den Gutachtenden gewählt. Dazu war es nötig, vom Pädagogischen Medienzentren Luzern

[14] http://edudoc.ch/record/96780/files/grundkomp_fremdsprachen_d.pdf (120103)
[15] MRECEF (2003 et 2009)

die CDs für das Hörverständnis zu erhalten (funktionierte bestens). Hilfreich wäre hier gewesen, auch die Korrekturschlüssel zur Verfügung zu haben.

3.4.4. Analyse der Teilfertigkeiten der Prüfungen

Hörverstehen
2009
Positiv sind die vielen einzelnen Hörtexte zu verschiedenen Themen. Dieser Prüfungsteil eignet sich gut für A2 (Alltag, kurze soz. Regelungen, einfacher direkter Austausch). Ein Hörverstehensteil auf B1.1 fehlt.
Die geringe Anzahl Items / Punkte (14) schränkte die Validität ein.

2010
17 Punkte sind wenig valide.
Dokument A und B: Items sind klar. Positiv ist auch das Vorhandensein von zwei unterschiedlichen Dokumenten.
Dokument A dürfte für das Prüfen des Niveaus A2.2 geeignet sein.
Dokument B eignet sich wahrscheinlich am besten für das Niveau B1 (Vokabular über Alltagsregelungen hinaus Richtung Interessen, Ereignisse. Der schnelle lang anhaltende Rhythmus des Texts und der Fragen lassen dieses B1 gegen den oberen Rand gehen (B1.2). Detail: Im einleitenden Text sollte die Kommunikationssituation erklärt werden. Wer ist Julie und Pierre, wer ist der Anrufer, wen ruft er an?

Fazite Hörverstehen 2009 und 2010
Aufgabenstellung und Fragen sind i. A. klar.
Die Punkteverteilung sollte auch pro einzelne Frage angegeben werden.
Die hauptsächlichen Itemformen Questions de Choix multiple QCM und Vrai/faux avec justification (welche rationelles Korrigieren ermöglichen) sind den KandidatInnen wohl vertraut. Sie sollten m. E. aber mit ein paar globaleren Fragen ergänzt werden. Warum? Es sollte auch geprüft werden, ob jemand das Thema und seine Hauptaussagen versteht. QCM und V/F kommen nicht darum herum, das Thema schon anzukündigen und Aussagen vorzuschlagen. Ähnliches kann auch zum Leseverstehen gesagt werden.
Die relativ geringe Anzahl Items / Punkte dieser beiden Hörverstehen wirken sich einschränkend auf die Validität aus.
Die Niveaus zwischen 2009 und 2010 sind unterschiedlich (von A2.1 bis gegen B1.2). Das Niveau B1.1 fehlt in 2009. Die Niveaus zum Hörverstehen sollten auf A2.2 bis B1.1 gehalten werden.

Leseverstehen

2009

Text A

Der Text A 2009 eignet sich, auf Grund der Sprachmittel (Spektrum „Freizeit" ist typisch B1), der Textarchitektur (Abfolge von Punkten ist typisch A2), des Texttyps (lineare Beschreibung ist typisch B1) gesamthaft für B1. Dies trotzdem die Themen eher nicht vertraut sein dürften (vertraute Themen sind ein Merkmal von A2 und B1).

Die Fragen zu Text A entlasten das B1 Richtung A2, weil einige Fragen Hinweise auf die relevante lokal begrenzte Textstelle enthalten[16] (identischer Wortlaut zwischen Frage und Text). Dieser Umstand wird etwas entschärft durch den relativ hohen Grad an Detailliertheit der Fragen.

Insgesamt ist der Fragekatalog möglicherweise aber nicht typisch B1 (B1 wären Fragen nach wesentlichen Aussagen, ohne klare Bezüge zu relevanten Textstellen).

Weitere Beobachtungen: Es sind zwei Fragetypen vorhanden (positiver Punkt). Die Frage g) respektiert nicht die Textaussage, wonach der Hang der Franzosen weg von der Arbeit hin zur Freizeit ein Eindruck (und nicht eine Tatsache) ist.

Text B

Das Vorhandensein eines zweiten Texts (Text B) ist begrüssenswert, da dies die Validität erhöht. Text B ist zudem ein anderer Texttyp als Text A. Leider ist er aber thematisch dem Text A recht nahe und trägt somit zur thematischen Repräsentativität des Niveaus wenig bei.

Text B eignet sich für B1.2 (Argumentation, soziale Fragen). Das Niveau B1.2 wird durch den Fragetyp Vrai / faux etwas gemildert, denn dieser gibt die Themen vor, so dass das eigenständige (globale) Suchen nach Botschaften / Argumenten / Hauptaussagen praktisch abwesend ist. Text B insgesamt dürfte sich also auf einem B1.1 bewegen.

2010

Merkmale dieser Lesetexte sind gutes B1.2 Niveau im Spektrum und in der Textarchitektur. Das Niveau B1.2 wird durch den Texttyp, welcher vertraut sein sollte und die nonverbale Entlastung durch die Bilder etwas gemildert. Die Fragen sind sprachlich praktisch auf einer B1.2 Schwierigkeit (anzustreben wäre A2). Insgesamt dürfte das Leseverstehen ein B1.2 repräsentieren.

Detail: Frage 4 erlaubt doch 2 Antworten, nicht?

Fazit Leseverstehen 2009 und 2010

Die Items (Fragen) sind i. A. klar, die gewählten Formate bekannt; diese sollten aber mit globaleren Formaten ergänzt werden.

Die Niveauabdeckung ist unterschiedlich gut. Im Gegensatz zu 2009 ist das Leseverstehen in 2010 enger (nur ein Texttyp, weniger breite Niveauabdeckung).

[16] siehe dazu Generalsekretariat der EDK – Projektleitung HarmoS Februar 2008, in www.oszhindelbank.ch/.../INFO-EDK-Standards-02-2008__2_.pdf

Es sollten immer zwei Lesetexte mit unterschiedlichen Themen, Texttypen, und Niveaus (A2.2 und B1.1) vorhanden sein. Das Niveau B1.2 sollte vermieden werden.
Die Sprache der „Questionnaire“ ist teilweise zu schwierig. Das Verstehen dieser Sprache dürfte einen hohen Anteil am Resultat haben, wahrscheinlich einen zu hohen. Abhelfen würden hier ein einfacheres Französisch und / oder Fragen (teilweise) in der Schulsprache (Deutsch)
Die Anzahl Punkte (17 bei 2010 und 20 bei 2009) ist zu wenig hoch. Angestrebt werden sollten min. 30.

3.4.5. Textproduktion

2009
Die Aufgabenstellung ist wenig kommunikativ (Informationen sind zu stark vorgegeben).
Der vorgeschriebene Umfang (60-80 Wörter) ist einiges zu gross für die vorgeschriebenen Aufgaben. 30-40 Wörter würden dafür genügen.

2010
Sie ist zu nahe an einer Übersetzung, zu weit weg von einer (auch im Lehrplan geforderten) Kommunikation, hinter welcher die einzelnen KandidatInnen selber stehen.
Der vorgeschriebene Umfang (60-80 Wörter) reicht zu knapp für die vorgeschriebenen Aufgaben.
Das Thema ist nahe dem Lesetext. Das ist nicht günstig, denn es schafft eine Abhängigkeit: wer den Lesetext gut bewältigt hat, profitiert hier, wer nicht, geht leer aus. Aber Lesen und Schreiben müssen unabhängig von einander geprüft werden.

Fazit Textproduktion 2009 und 2010
Textproduktionen sind bezüglich der Themen und der Texttypen lehrplankonform. Durch die ausführlichen inhaltlichen Vorgaben sind sie jedoch zu nahe an der Übersetzung, resp. zu weit weg von echter Kommunikation und haben deshalb eine eingeschränkte Konstruktvalidität.
Bei der Aufgabenstellung ist besser zu kontrollieren, dass die verlangte Anzahl Wörter mit der Aufgabe kompatibel ist.
Die thematisch-lexikalische Unabhängigkeit der Textproduktion vom Hör- und Leseverstehen ist besser zu beachten.
Zwei Texte schreiben lassen (mit unterschiedlichen Themen, Textsorten) statt einen wäre valider.
Es fehlen die Angaben über Kriterien (für mich als Gutachter, aber auch für die Kandidatinnen), und diese müssen detailliert und mit den Punkten verbunden sein. Es muss klar deklariert werden, welche Qualitäten beobachtet werden und was innerhalb dieser Qualität wie viel wert ist (sehr gut, gut genügend, ungenügend,

...). Eine Ankoppelung der Evaluation mit dem gemeinsamen europäischen Referenzrahmen oder allenfalls Lingualevel wäre empfehlenswert[17].

3.4.6. Analyse allgemeiner Eigenschaften der Prüfungen

Information zuhanden der KandidatInnen
Wünschenswert wäre, zum Beispiel auf der Titelseite, eine Information in Form von kurzen Beispieldeskriptoren aus dem Lehrplan (mit der GER und dem Europäischen Sprachenportfolio ESP im Hintergrund) an die KandidatInnen darüber, welche Teilfertigkeiten auf welchem GER-Niveau geprüft werden sollen.

Fehlen der Teilfertigkeiten „Sprechen"
Es fehlen die mündliche Interaktion und Produktion. Diese Teilfertigkeiten sind im Lehrplan wichtig. Sie nicht zu prüfen, könnte negative Folgen (washback) für den Unterricht in der Sek I haben (Vernachlässigung). Auf jeden Fall ist es ein den offiziellen Bildungs- und Ausbildungszielen des Fremdsprachenunterrichts entgegen laufende Praxis. Es sollte deshalb eine praktikable Möglichkeit gefunden werden, diese Fertigkeiten zu prüfen. Falls eine flächendeckende Prüfung angesichts der Diskrepanz zwischen Anzahl KandidatInnen und Anzahl möglicher Ausbildungsplätze nicht in Frage kommt, müssen m. E. andere Lösungen angestrebt werden. Eine Zusammenarbeit mit den abgebenden Schulen wäre hier eventuell eine Strategie, eine andere wäre es vielleicht KandidatInnen in einer zweiten Phase mündlich zu prüfen, welche in den andern schriftlich prüfbaren Teilfertigkeiten nur knapp genügend oder ungenügend sind.

Gewichtung der Teilfertigkeiten
Gemäss Lehrplan müsste Hörverstehen am meisten gewichtet werden, Schreiben am wenigsten. Dies ist in der Prüfung 2010 nicht ganz der Fall, in der Prüfung 2009 klar nicht (siehe Punkteverteilung). Die Erhöhung der Gewichtung des Hörverstehens könnte mit der auch nötigen Erhöhung der Punktezahl gekoppelt werden.

Niveaus / Schwierigkeiten der Prüfungsteile
Es ist anzunehmen, dass das Gros der Kandidaten/innen in wirklich selbständigen, kommunikativen Angaben ein A2.2, einige vielleicht ein B1.1 aufweisen. Allgemein bewegt sich die Prüfung diesbezüglich auf einem validen Niveau. Allerdings sollten B1.2 Aufgaben wegfallen.
Die Prüfungsteile und –jahrgänge sind unterschiedlich schwierig. So besteht ein erheblicher Unterschied zum Beispiel zwischen Hörverstehen 2009 und 2010. Diese Unterschiede sollte besser kontrolliert werden. Es ist auch nicht ersichtlich, ob diese Unterschiede in der Beurteilung berücksichtig werden.

[17] www.lingualevel.ch/

Sprache der Aufgabenstellungen
Wie schon in der Diskussion des Hör- und Leseverständnisses erwähnt, sollen die Sprache und die Formulierung der Aufgabenstellungen auf keinen Fall den Zugang zur eigentlichen Aufgabe erschweren. Wenn die Aufgabenstellung in Französisch erfolgt, soll sie klar auf Niveau A2 sein, oder, wenn nicht, durch Hilfen wie Übersetzungen entlastet werden. Und warum nicht, wie das in der Textproduktion schon gemacht wird, oder zum Beispiel bei Lingualevel, die Aufgaben auf Deutsch (Schulsprache) stellen.

Kriterien / Notengebung
Nochmals erwähnt sei hier das Fehlen der Beurteilungskriterien für die Textproduktion. Zudem steht auf den Prüfungsblättern nichts zu den Berechnungsgrundlagen der Noten. Diese Information sollten die KandidatInnen auf dem Blatt sehen.

Gebrauch Wörterbuch
Gebrauch eines Wörterbuchs ist in den Niveaus bis B2 nicht üblich, monolinguale ab B2 / C1 für bestimmte Fertigkeiten[18]. Zweisprachige Dictionnaires dürften auch in der Sek I bei Prüfungen nicht erlaubt sein. Der Verzicht darauf würde die Ausgangslage für alle KandidatInnen einheitlicher machen.

3.4.7. Zusammenfassung

Die Prüfung ist solide und sorgfältig gestaltet. Sie prüft effektiv Hör- und Leseverstehen, in der Textproduktion sollte der kommunikative Anteil erhöht und Transparenz bezüglich der Kriterien geschaffen werden. Die Prüfung scheint mir unsicher was die Kontrolle ihrer Schwierigkeit, resp. ihres Niveaus anbelangt. Grob liegt sie zwar richtig, im Einzelnen sind die Unterschiede jedoch zu gross. Ein stringenteres, an den Lehrplan (und damit an den GER und das ESP) sich anlehnendes konzeptuelles Vorgehen beim Erstellen der Prüfung scheint mir nötig. In allen drei Teilfertigkeiten sollte die Prüfung quantitativ noch wachsen, besonders im Hörverständnis und in der Textproduktion. Diese Befunde sowie das Nicht-Prüfen des Sprechens schränken die Validität der Prüfung ein.

[18] siehe www.delfdalf.ch

4. Vorhersage des BM-Abschlusses (Modul 2)

In diesem Modul wurde die prognostische Validität der Noten der Aufnahmeprüfung untersucht und damit die Frage, ob und ggf. wie stark der Erfolg bei der Aufnahmeprüfung mit dem Erfolg bei der Abschlussprüfung zur Erlangung der Berufs- oder Fachmaturität zusammenhängt. Zu diesem Zweck wurden die fachbezogenen Noten beider Prüfungen miteinander korreliert. Neben den Korrelationen zwischen den gleichen Fächern wurden auch Zusammenhänge zwischen verschiedenen Fächern berechnet (vgl. Tabelle 11).

Tabelle 11: Korrelationsmatrix Noten Aufnahme- und Abschlussprüfung

r_s	AP_AA	AP_DE	AP_FR	AP_EN	S_AA	S_DE	S_FR	S_EN	S_⌀
AP_AA	1.000	-.06	-.03	-.02	**.42****	-.05	-.01	.06	.23**
AP_DE		1.000	.10	.14**	-.02	**.31****	.14*	.10	.14*
AP_FR			1.000	.34**	.03	.20**	**.56****	.29**	.22**
AP_EN				1.000	-.03	.25**	.24**	**.57****	.17**
S_AA					1.000	.14**	.18**	.03	.69**
S_DE						1.000	.37**	.37**	.54**
S_FR							1.000	.39**	.50**
S_EN								1.000	.37**
S_⌀									1.000

Legende

$300 \leq n \leq 458$

* Die Korrelation ist auf dem 5%-Niveau signifikant (zweiseitig).

** Die Korrelation ist auf dem 1%-Niveau signifikant (zweiseitig).

AP = Aufnahmeprüfung, S = Abschlussprüfung, AA = Arithmetik und Algebra; DE = Deutsch, FR = Französisch, EN = Englisch, ⌀ = Durchschnittsnote

Die Noten der Aufnahmeprüfung korrelieren im gleichen Fach zwischen $r_s = .31$ (Deutsch) und $r_s = .57$ (Englisch) mit den entsprechenden Abschlussnoten. Diese moderaten bis hohen Zusammenhänge liegen im erwarteten Bereich. Interessant ist im Übrigen, dass die Mathematiknote völlig unabhängig ist von den Noten in den Sprachfächern, während die Leistungen in den beiden Fremdsprachen sowohl im Querschnitt als auch im Längsschnitt deutlich zusammenhängen.

5. Vergleich der beiden Zulassungsvarianten (Modul 3)

In diesem Modul wurden die Schüler/innen, die prüfungsfrei zugelassen worden waren, mit den Schüler/-innen verglichen, die die Aufnahmeprüfung bestanden hatten. Im Zentrum stand dabei die Frage, ob das Kriterium für den prüfungsfreien Zugang (Notendurchschnitt von mind. 4.7 für Sek-A) angemessenen ist oder nicht. Mögliche Indikatoren sind hier Schulnoten, provisorische Promovierungen und Drop outs. Als Stichprobe wurden die im Jahr 2010 zugelassenen Lernenden untersucht. Grundlage für den Vergleich waren die Indikatoren ein Jahr nach dem Eintritt.

5.1. Aufnahmeverfahren, Promovierungen und Dropout-Raten

Ob zwischen dem Aufnahmeverfahren und der Promovierung ein Jahr nach dem Eintritt ein Zusammenhang besteht, wurde statistisch mittels Chi-Quadrat getestet (vgl. Tabelle 12).

Tabelle 12: Aufnahmeverfahren, Promovierungen und Drop outs

	Aufnahmeverfahren				
	mit Prüfung (n=192) %	ohne Prüfung (n=161) %	χ^2	df	p
provisorisch promoviert	15	9	3.61	1	0.057
Ausbildung abgebrochen	6	6	0.04	1	0.849

Der Anteil der Schülerinnen und Schüler, die provisorisch promoviert wurden, beträgt bei den über eine Prüfung aufgenommenen Lernenden 15% und ist somit höher als bei den prüfungsfrei aufgenommenen Lernenden (9%). Dieser Unterschied ist statistisch gesehen tendenziell signifikant. Demgegenüber unterscheiden sich die beiden Aufnahmevarianten bezüglich der Drop out-Raten (zweimal 6%) nicht.

5.2. Aufnahmeverfahren und Noten

Um zu untersuchen, ob sich prüfungsfrei Aufgenommene von den übrigen Schüler und Schülerinnen nach einem Jahr bezüglich der erhaltenen Noten unterscheiden, wurden die Notenmittelwerte der beiden Gruppen mittels t-Tests verglichen. Die ohne Aufnahmeprüfung an die BMS zugelassenen Schülerinnen und Schüler hatten nach einem Jahr signifikant bessere Noten in Arithmetik / Algebra, Französisch und Englisch als die via Prüfung aufgenommenen Schülerinnen und Schüler (vgl. Tabelle 13). Die betreffenden Mittelwerte sind signifikant unterschiedlich. Keine Unterschiede fanden wir in Bezug auf die Leistungen im Fach Deutsch.

Tabelle 13: Mittelwertvergleich Noten mit / ohne AP

	AP	N	Mittelwert	SD	t	df	p
Note_AA	mit	178	4.18	.877	-2.78	330	.006
	ohne	154	4.44	.802			
Note_DE	mit	178	4.573	.502	-1.47	330	.142
	ohne	154	4.653	.479			
Note_FR	mit	176	4.315	.692	-3.31	329	.001
	ohne	155	4.558	.637			
Note_EN	mit	146	4.55	.599	-3.21	257	.001
	ohne	113	4.79	.550			

6. Analyse der Kompetenzniveaus (Modul 4)

In diesem Abschnitt werden die Ergebnisse der Raschskalierung der Prüfungen in Arithmetik / Algebra und Deutsch für die Jahre 2009 und 2010 berichtet. Die Darstellung der Personen- und Itemverteilung[19] einer Prüfung erfolgt in Form einer Graphik. Die daran anschliessende Tabelle zeigt jeweils die durchschnittlichen Schwierigkeiten der Aufgaben.
Zum Schluss werden noch das DIF, das bei zwei Aufgaben aufgetreten ist und Probleme bei Vergabe von halben Punkten besprochen.

6.1. Raschskalierung AP Arithmetik und Algebra 2009 und 2010

Die Verteilung der Schüler/innen auf die Arithmetik/Algebra-Aufgaben 2009 findet sich in Abbildung 5.

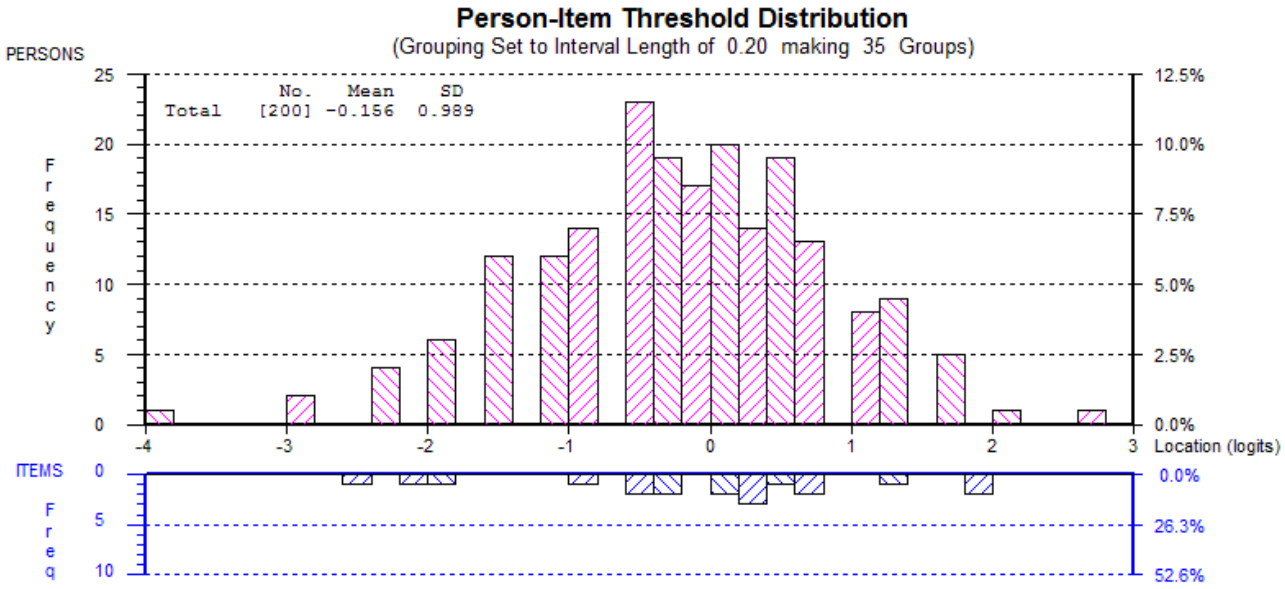

Abbildung 5: Personen- / Itemverteilung AP Arithmetik und Algebra 2009

In Tabelle 14 sind die Prüfungsaufgaben der AP Arithmetik und Algebra 2009 entsprechend ihrer Schwierigkeit (von schwierig bis leicht) angeordnet. Da

[19] Item = Aufgabe. Die Grafiken zeigen nicht die Verteilung der Aufgaben, sondern die Verteilung der Thresholds (Schwellenwerte) der Aufgaben. Die Prüfungsaufgaben haben nicht nur zwei z.B. „falsch“ oder „richtig“, sondern mehrere Stufen z.B. falsch, teilweise richtig, richtig als Lösung. Bei der Raschskalierung werden Thresholds für die einzelnen Aufgaben berechnet. Eine Aufgabe lässt sich abstufen aufgrund dieser Schwellenwerte. Einige Examinanden werden keine, einen Teil oder alle Schwellen einer Aufgabe überschreiten.

Aufgabe 6 ein erhöhtes DIF[20] bezüglich des Alters der Prüfungsteilnehmer/innen aufweist, wurde ihre Schwierigkeit für drei Altersgruppen (unter 16 Jahre, 17-18 Jahr, über 18 Jahre) separat berechnet.

Tabelle 14: Mittlere Aufgabenschwierigkeit AP AA2009

Aufgabe	Schwierigkeit	Aufgabe	Schwierigkeit
Aufgabe 6.1	1.409	Aufgabe 8	-0.243
Aufgabe 5	1.014	Aufgabe 1	-0.525
Aufgabe 6.2	0.889	Aufgabe 2	-0.532
Aufgabe 6.3	0.023	Aufgabe 4	-0.538
Aufgabe 3	-0.002	Aufgabe 7	-1.495

Die durchschnittliche Personenfähigkeit betrug -0.3497. Die meisten Aufgaben (Items) befinden sich in Abbildung 5 oberhalb der durchschnittlichen Personenfähigkeit. Sie sind also insgesamt etwas zu schwierig. Mit einem Wert von 0.6979 war der Separationsindex[21] gerade genügend gross, um eine Raschskalierung durchzuführen.

Das Ergebnis der Raschskalierung der Arithmetik/Algebra-Prüfung 2010 ist aus der Abbildung 6 ersichtlich. Wie man in dieser Grafik sieht, sind die meisten Aufgaben in der Mitte anzutreffen, wo sich viele Schüler mit einer mittleren Fähigkeit befinden. Mit einem Wert von 0.7653 war der Separationsindex gross genug, um die Raschskalierung durchzuführen.

In Tabelle 15 sind die Prüfungsaufgaben der AP Arithmetik und Algebra 2010 entsprechend ihrer Schwierigkeit (von schwierig bis leicht) aufgeführt. Da Aufgabe 8 ein DIF bezüglich des Geschlechts der Kandidaten/innen aufwies, wurde ihre Schwierigkeit für Mädchen und Jungen separat berechnet.

[20] DIF = Differential Item Functioning. Ein erhöhter DIF-Wert bedeutet, dass Menschen aus verschiedenen Gruppen (z.B. Geschlecht, Alter, Ethnie) mit derselben Fähigkeit eine unterschiedliche Wahrscheinlichkeit aufweisen, eine bestimmte Aufgabe zu lösen (vgl. z.B. http://de.wikipedia.org/wiki/Differential_item_functioning, 19.4.12)

[21] Der Separationsindex sagt etwas darüber aus, wie gut Personen in verschiedene Fähigkeitsgruppen unterteilt werden können. Er sollte mindestens bei 0.7 liegen, damit sich zwei Gruppen, z.B. schlecht und gut, bilden oder unterscheiden lassen. Die Modellberechnung der AP AA2010 ergab z.B. einen Separationsindex von 0.76 d.h. 76 % der Varianz werden durch das Raschmodell erklärt.

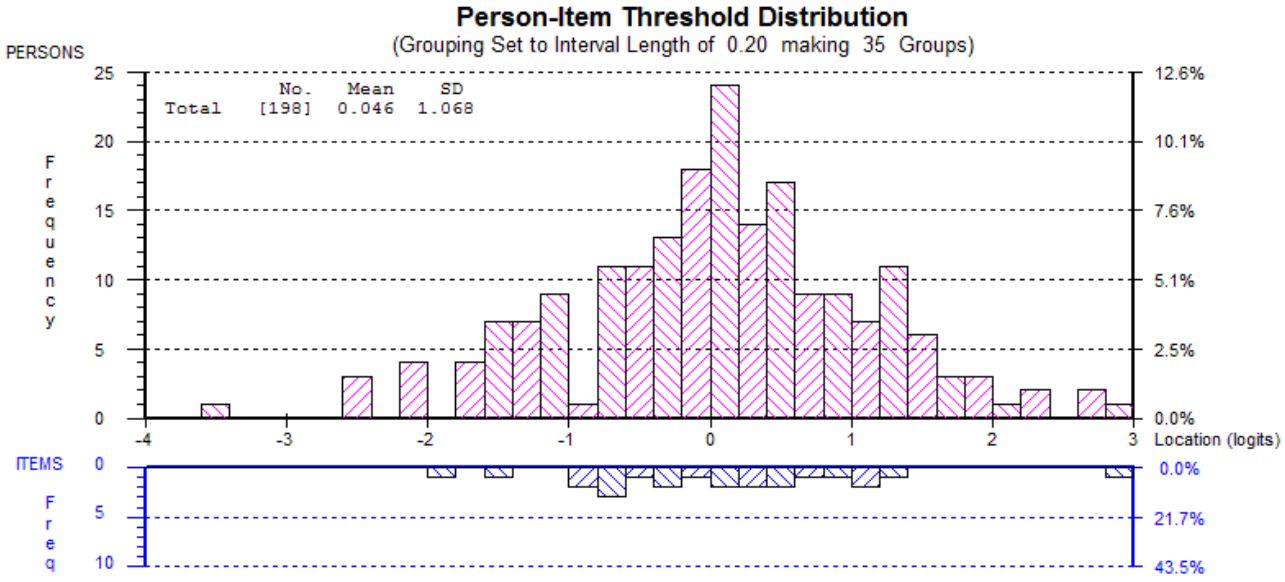

Abbildung 6: Personen- / Itemverteilung, AP Arithmetik und Algebra 2010

Tabelle 15: Mittlere Aufgabenschwierigkeit AP AA 2010

Aufgabe	Schwierigkeit	Aufgabe	Schwierigkeit
Aufgabe 5	1.136	Aufgabe 3	-0.497
Aufgabe 4	0.891	Aufgabe 8m	-0.535
Aufgabe 2	0.402	Aufgabe 7	-0.678
Aufgabe 8w	0.259	Aufgabe 1	-0.756
Aufgabe 6	-0.223		

DIF (Differential Item Functioning)

Bei zwei Aufgaben zeigte sich ein sogenanntes DIF (s.o.), d.h. dass Menschen aus verschiedenen Gruppen (z.B. Geschlecht, Alter, Ethnie) mit derselben Fähigkeit eine unterschiedliche Wahrscheinlichkeit aufweisen, die Aufgabe zu lösen. Aufgabe 6 der AP 2009 (vgl. Abbildung 7) diskriminiert nach dem Alter. Es handelt sich um eine Zinsrechnung, welche Personen, die älter als 18 Jahre sind, besser lösen, als jüngere (<16 Jahre).

Möglicherweise spielt beim Lösen dieser Aufgabe die Lebenserfahrung eine gewisse Rolle: Ältere Personen mussten sich in ihrem Alltag bereits vermehrt mit Zinsrechnung befassen als jüngere. Es sollte aber nicht so sein, dass eine Aufgabe nach dem Alter diskriminiert.

6. Eine Person möchte sich einmal für einen Tag als Millionär fühlen.

a) Wie viele Franken würde sie dieser Spass kosten, wenn sie annimmt, dass die Bank dafür einen Zins von 7,8 % verlangt. Runden Sie auf 5 Rp. genau (und rechnen Sie in der Teilaufgabe b) mit dem gerundeten Resultat weiter).

b) Wie viele Tage müsste die Person den Betrag von Fr. 12'415.– auf ihrem Sparkonto zu 2,25 % an Zins legen, um sich mit dem erhaltenen Zins dieses Gefühl leisten zu können? (Auf ganze Tage aufrunden)

Abbildung 7: Aufgabe 6, AP 2009

Um die Unterschiede aufzuzeigen, wurden drei Gruppen gebildet (≤16-Jährige, 17- bis 18-Jährige und >18-Jährige). Bei der Gruppe >18-Jährigen zeigte sich ein durchschnittlicher Schwierigkeitsgrad für diese Aufgabe (vgl. Abbildung 8). Bei der Gruppe ≤16 zeigte sich ein hoher Schwierigkeitsgrad.

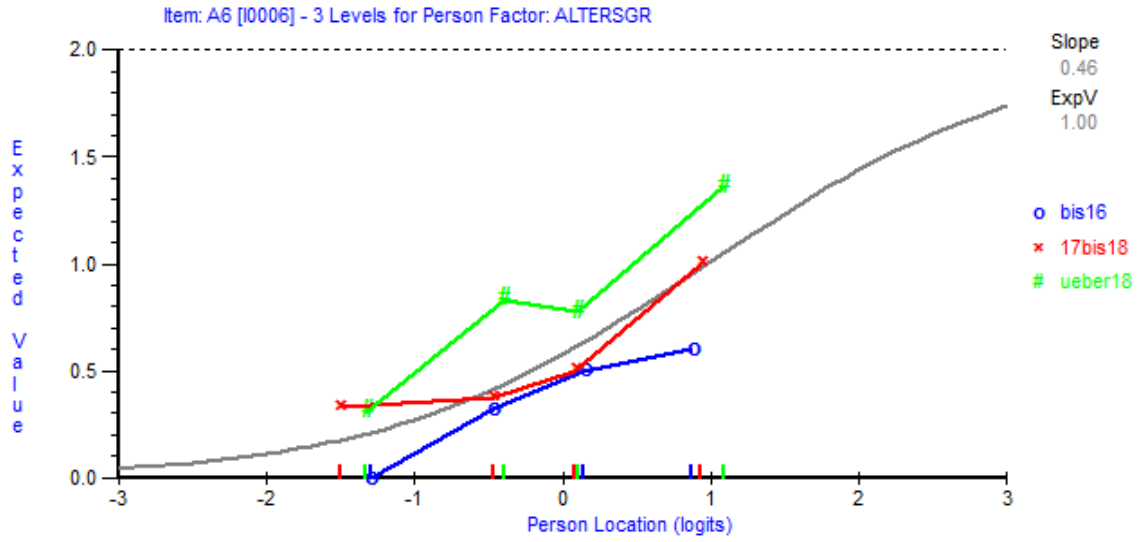

Abbildung 8: Abbildung 6: Aufgabe 6, ICC (Item characteristic curve) separiert nach Altersgruppen

Bei Aufgabe 8, AP 2010 (Abbildung 9) der AP 2010 wurde ein DIF in Bezug auf das Geschlecht festgestellt. Aus Abbildung 10 geht hervor, dass die Aufgabe von Jungen signifikant besser gelöst wurde als von Mädchen.

8. Wenn Marco mit dem Bike jeweils um 07.10 h abfährt, kommt er um 07.24 h im 4.9 km entfernten Lehrbetrieb an. An einem Morgen trifft er nach 5 Minuten einen Kollegen und versäumt daher 2 Minuten.
Mit welcher Geschwindigkeit in km/h muss er nun weiterfahren, wenn er zur üblichen Zeit bei seinem Lehrbetrieb eintreffen will?

Abbildung 9: Aufgabe 8, AP 2010

Werden bei der Berechnung des Raschmodells die Geschlechter nicht unterschieden, ergibt sich für die Aufgabe 8 eine Schwierigkeit von -0.085. Die separate Berechnung ergibt für Jungen eine Schwierigkeit von -0.5, für die Mädchen 0.259 (die Schwierigkeiten liegen im Normalfall zwischen -2 und +2). Die Aufgabe 8 ist somit für Jungen einfach, aber überdurchschnittlich schwierig für Mädchen.
Eine Erklärung könnte in den geschlechtsspezifisch variierenden alltäglichen Lebenswelten liegen. Items sollten aber für alle demographischen Gruppen (z.B. Alter, Geschlecht, Nationalität usw.) gleich funktionieren.
Um dieses DIF auszugleichen, könnte die Aufgabe abhängig vom Geschlecht unterschiedlich bewertet werden, d.h. Mädchen erhielten mehr Punkte. Da dies für eine Aufnahmeprüfung nicht realistisch ist, könnte eine andere Lösung darin bestehen, als Ausgleich eine Aufgabe anzubieten, die sich vor allem auf den Kontext von Mädchen bezieht.

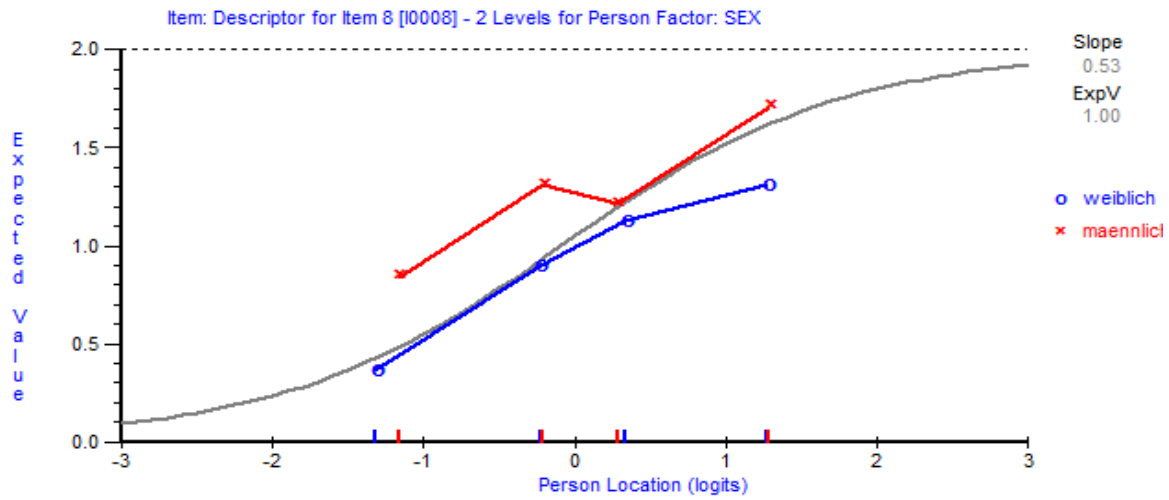

Abbildung 10: Aufgabe 8, ICC (Item characteristic curve) separiert nach Geschlecht

6.2. Raschskalierung AP Deutsch 2009 und 2010

Das Ergebnis der Raschskalierung der AP Deutsch 2009 ist in Abbildung 11 dargestellt.
Die durchschnittliche Personenfähigkeit betrug 0.445. Die meisten Aufgaben befinden sich unterhalb der durchschnittlichen Personenfähigkeit, was bedeutet, dass die Prüfung etwas zu viele einfache Aufgaben aufwies. Die geringe Standardabweichung (0.52) der Personenfähigkeit weist ausserdem auf eine geringe Streuung der Aufgaben hin.
Bei einem Separationsindex von 0.617 war die Prüfung nur knapp raschskalierbar: Die Prüfung diskriminierte zu wenig gut zwischen verschiedenen Personen und war deshalb nur begrenzt geeignet bzw. suboptimal um Deutschkompetenzen zu messen.[22]

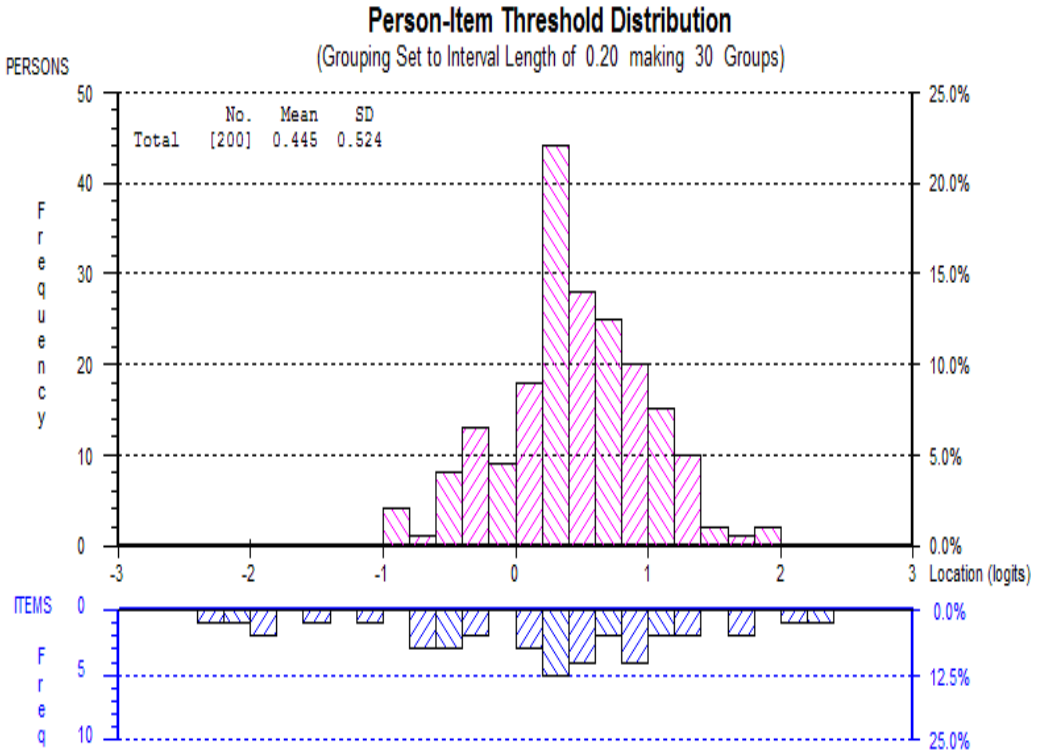

Abbildung 11: Personen- / Itemverteilung AP Deutsch 2009

[22] Das ist hier vielleicht aber auch gar nicht gefragt. Der Test kann trotzdem dazu dienen, einen Cutoff bzw. eine Grenze zwischen "bestanden" und "nicht bestanden" zu bestimmen.

Tabelle 16: mittlere Aufgabenschwierigkeit AP Deutsch 2009

Aufgabe	Schwierigkeit	Aufgabe	Schwierigkeit
Aufgabe 13	1.061	Aufgabe 12	0.046
Aufgabe 14	0.831	Aufgabe 4	0.01
Aufgabe 1	0.542	Aufgabe 10	-0.068
Aufgabe 15	0.457	Aufgabe 3	-0.214
Aufgabe 16	0.383	Aufgabe 8	-0.41
Aufgabe 9	0.346	Aufgabe 5	-0.527
Aufgabe 11	0.317	Aufgabe 6	-0.931
Aufgabe 7	0.176	Aufgabe 2	-2.017

In Tabelle 16 sind die Prüfungsaufgaben der AP Deutsch 2009 entsprechend ihrer Schwierigkeit (von schwierig bis leicht) angeordnet.

Das Ergebnis der Raschskalierung der AP Deutsch 2010 ist in Abbildung 12 dargestellt. Die durchschnittliche Personenfähigkeit beträgt 0.255. Die Prüfung war somit insgesamt eher zu leicht, bei einem Übergewicht an Items mit mittlerer Schwierigkeit. Die Standardabweichung der Personenfähigkeit betrug 0.62, was eher gering ist.

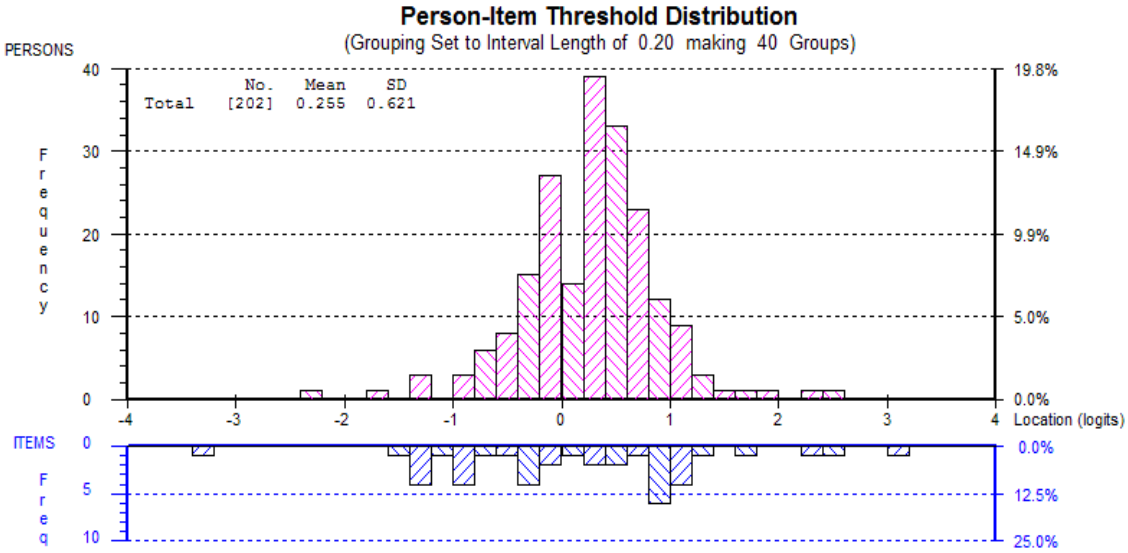

Abbildung 12: Personen- / Itemverteilung AP Deutsch 2010

Mit einem Separationsindex von 0.679 war die Prüfung nur knapp raschskalierbar und es ergab sich keine gute Skala. Auch diese Prüfung diskriminierte zu wenig

gut zwischen verschiedenen Personen und war deshalb auch nicht geeignet bzw. suboptimal um Deutschkompetenzen zu messen.

In Tabelle 17 sind die Prüfungsaufgaben der AP Deutsch 2010 entsprechend ihrer Schwierigkeit (von schwierig bis leicht) angeordnet.

Tabelle 17: Mittlere Aufgabenschwierigkeit AP Deutsch 2010

Aufgabe	Schwierigkeit	Aufgabe	Schwierigkeit
I0016	2.264	I0011	-0.028
I0009	1.262	I0004	-0.116
I0022	1.104	I0008	-0.17
I0005	0.873	I0015	-0.177
I0018	0.846	I0017	-0.265
I0010	0.691	I0021	-0.597
I0020	0.489	I0012	-0.772
I0013	0.382	I0007	-0.894
I0003	0.346	I0001	-0.968
I0002	0.077	I0014	-1.114
I0019	0.068	I0006	-3.3

7. Fazit

1. Die Prüfungsaufgaben sind im Allgemeinen inhaltlich valide. Sie prüfen Inhalte, die mit dem Lehrplan der Sekundarstufe I übereinstimmen.
2. Einzelne Inhalte der Lehrpläne fehlen (z.B. Hören und Sprechen in den Bereichen Deutsch oder Französisch), was allerdings aufgrund des damit verbundenen Prüfungsaufwands nachvollziehbar ist.
3. Einzelne Aufgaben der Mathematikprüfung sind – gemessen am Lehrplan – eher etwas zu schwierig. Bei den Korrekturlösungen könnte die zeitgemässe didaktische Forderung nach multiplen Lösungswegen vermehrt eingelöst werden.
4. Die Gewichtung der Inhalte (innerhalb der gleichen Prüfung) ist im Allgemeinen richtig. Bei einzelnen Gewichtungen sollten allerdings Alternativen erwogen werden (z.B. stärkere Gewichtung des Sprachhandelns im Fach Deutsch).
5. Einzelne Prüfungsteile könnten noch wachsen. Dies würde ihre Reliabilität erhöhen (z.B. Hörverständnis Französisch).
6. Die Korrelationen zwischen den Noten der Aufnahmeprüfung und der Schlussprüfung im gleichen Fach sind statistisch bedeutsam. Die Aufnahmeprüfung weist somit eine gute prognostische Validität auf.
7. Ohne Prüfung (aufgrund des Notendurchschnitts der abgebenden Schulen) aufgenommene Schüler/innen werden etwas seltener provisorisch promoviert als Schüler/innen, die aufgrund einer bestandenen Aufnahmeprüfung aufgenommen wurden. Erstere haben in verschiedenen Fächern auch die besseren Noten. Dies spricht für die Beibehaltung der gegenwärtigen Aufnahmepraxis (bzgl. Notendurchschnitt für prüfungsfreie Zulassung).
8. Die Mathematikprüfung enthält unterschiedlich schwierige Aufgaben und kann dadurch unterschiedliche Kompetenzniveaus messen.
9. Die Deutschprüfung deckt ein deutlich weniger breites Spektrum an Kompetenzen ab. Die zusätzliche Verwendung von anspruchsvolleren Aufgaben würde hier Abhilfe schaffen.

8. Literatur

Bühner, M. (2006). Einführung in die Test- und Fragebogenkonstruktion. München: Pearson Studium.

Hodel, H-P. (2007). Vergleichbares Prüfen lokal umsetzen. In Babylonia 1/2007.

Milanovic (A.L.T.E.), M. (dir.): Evaluation de compétences en langues et conception de tests. Cadre européen commun de référence pour les langues. Apprendre, Enseigner, Evaluer. Division des Politiques Linguistiques. Strasbourg, octobre 2002 (online)

MRECEF (2003 et 2009) Manual for relating Language Examination to CEF; Langague examining and testing development http://www.coe.int/T/DG4/Portfolio/?L=E&M=/documents_intro/Manual.html http://www.coe.int/t/dg4/linguistic/Manuel1_EN.asp) http://www.coe.int/t/dg4/linguistic/manuel1_FR.asp?

North, B. (2004). Relating Assessments, Examinations and Courses to the CEF. In K. Morrow (Ed.) Insights from the Common European Framework. Oxford: Oxford University Press.

Schneider, G. & North, B. (2000). Fremdsprachen können – was heisst das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit. NFP 33. Chur / Zürich: Verlag Rüegger.

Strobl, C. (2010). Das Rasch-Modell. Eine verständliche Einführung für Studium und Praxis. Bd. 2: Sozialwissenschaftliche Forschungsmethoden. München: Rainer Hampp.

Taglante, C. (2005) L'évaluation et le Cadre européen commun, Paris, CLÉ International.

Weir, C. J. (2005). Language Testing and Validation. Palgrave Macmillan.

Wilson, M. (2005). Constructing Measures: An Item Response Modeling Approach. New Jersey. Lawrence Erlbaum Ass. Publishers.