

Formative assessment im Grammatikunterricht für angehende Grundschullehrkräfte

Jana Brunner, Humboldt-Universität zu Berlin

Dr. Jana Brunner unterrichtet Linguistik an der Humboldt-Universität zu Berlin. Ihre Forschungsinteressen liegen im Bereich der Graphematik und Phonetik.

Email: jana.brunner@hu-berlin.de

Abstract: Lernplattformen wie Moodle bieten die Möglichkeit, kursbegleitende Übungsmodule zu erstellen. In diesem Artikel wird ein solches Modul für die Hochschullehre vorgestellt und evaluiert. Es wird gezeigt, dass das Absolvieren des Moduls einen stärkeren Einfluss auf die Abschlussnote hat als andere wichtige Einflussfaktoren wie beispielsweise die Teilnahme am Seminar.

Schlüsselwörter: formative assessment, summative assessment, blended learning, Moodle, Grammatikunterricht

Danksagung: Ich danke meiner Kollegin Eva Meier für wertvolle Kommentare zu einer früheren Version dieses Artikels.

1. Einleitung

Die Ausbildung von Studierenden an Hochschulen ist heute ohne Lernplattformen wie beispielsweise Moodle kaum noch denkbar. Während diese Plattformen vor 20 Jahren noch vor allem dazu genutzt wurden, Materialien zur Verfügung zu stellen, die vorher in Papierform an die Studierenden ausgeteilt wurden, werden die Plattformen heute immer stärker direkt für das Lernen genutzt. Beispielsweise kann man mittels einer solchen Plattform interaktive Übungen wie etwa Vokabelquizzes erstellen. Auch Prüfungen werden immer öfter digital absolviert. Das hat verschiedene organisatorische Vorteile: Abgaben werden genau registriert, studentische Arbeiten können nicht verloren gehen und sie können – mit den notwendigen Berechtigungen – von überall auf der Welt eingesehen werden. Wenn es der Lerngegenstand zulässt, eine Menge von richtigen Antworten vorab zu definieren, haben digitale Prüfungen einen weiteren Vorteil: Die Bewertung kann automatisiert erfolgen. Im Folgenden wird ein in Moodle implementiertes Modul beschrieben, welches diese Vorteile digitaler Tests nutzt. Das Modul wird kursbegleitend in der linguistischen Ausbildung von angehenden Grundschullehrkräften verwendet. Zunächst wird dargelegt, welche Überlegungen hinter der Struktur des kursbegleitenden Moduls stehen. Dann wird das Modul selbst beschrieben und evaluiert.

Bei Tests, egal ob digital oder nicht, kann man zwischen lernbegleitenden Tests (*formative assessment*) und Abschlussprüfungen (*summative assessment*) unterscheiden. Während beide Formate letzten Endes die Leistungen von Lernenden bewerten, haben sie darüber hinaus sehr unterschiedliche Ziele. Abschlussprüfungen (summative Tests) haben im Großen und Ganzen zwei Funktionen:

- Sie sollen darüber Aufschluss geben, ob Lernende über soviel Wissen zu einem Thema verfügen, dass sie die Lerneinheit oder einen Bildungsgang abschließen können.
- Erfolgt eine Benotung, die über die Unterscheidung zwischen „bestanden“ und „nicht bestanden“ hinausgeht, haben Abschlussprüfungen außerdem die Funktion, Qualifikationsunterschiede innerhalb der Gruppe, die bestanden hat, sichtbar zu machen.

Summative Tests sind für die genannten Zwecke sinnvoll, sie haben aber eine Reihe von Nachteilen. In der Regel sind sie ungeeignet, Lernenden Rückmeldungen zu geben, aus denen sie fachlich etwas lernen können. Summative Tests können Lernende auch kaum motivieren, sich weiterhin mit dem Thema zu beschäftigen, um mehr zu lernen, denn für die Lernenden ist das Thema mit dem Test abgeschlossen. Für diesen letztgenannten Zweck sind formative Tests (lernbegleitende Tests) jedoch sehr gut geeignet. Formative Tests haben eine ähnliche Funktion wie korrigierte Übungen oder Hausaufgaben. Sie finden während des Lernprozesses (zum Beispiel täglich oder wöchentlich) statt. Formative Tests haben also ganz andere Ziele als summative Tests:

- Die Lernenden sollen beim Absolvieren der Aufgaben fachlich etwas lernen (beim summativen Testen sollen sie nur zeigen, dass sie etwas können).
- Die Lernenden sollen eine Rückmeldung bekommen, mit deren Hilfe sie sich fachliches Wissen aneignen können.
- Die Lernenden sollen Empfehlungen bekommen, welches die nächsten Schritte in ihrem Lernprozess sein sollten.
- Die Lernenden sollen motiviert werden, sich möglichst intensiv mit dem Lerngegenstand zu befassen.

Lehrende gestalten summative und formative Tests häufig sehr ähnlich. Das ist nicht ideal, denn so können die unterschiedlichen Ziele dieser Testverfahren nicht erreicht werden. Beispielsweise erhalten Lernende bei einem formativen Test oft einfach nur eine Note als Rückmeldung. Um durch den Test etwas zu lernen, fehlen unter Umständen Informationen dazu, was eine richtige Antwort gewesen wäre und warum die von den Lernenden gegebenen Antworten falsch sind. Da summative und formative Tests aber unterschiedliche Ziele haben, sollten die Tests auch unterschiedlich angelegt sein, was (1) den Schwierigkeitsgrad angeht, (2) die Rückmeldung, die die Lernenden bekommen, (3) die Möglichkeit der Wiederholung des Tests.

Ziel eines summativen Tests ist es, zu erfahren, wieviel die Lernenden wissen oder können. Man gestaltet summative Tests deshalb am besten so, dass Testaufgaben enthalten sind, die viele Lernende unterfordern, aber auch einige, die die meisten Lernenden überfordern, denn nur so kann man erreichen, dass Lernende mit höherer Kompetenz eine bessere Bewertung bekommen als solche mit geringerer Kompetenz.

Bei formativen Tests sollte das Niveau an das Wissen der einzelnen Lernenden angepasst sein. Lernende mit viel Vorwissen sollten nicht dieselben Tests absolvieren wie Lernende mit wenigen Vorkenntnissen. Sie sollten einerseits keine Zeit mit der Bearbeitung von zu leichten Aufgaben vergeuden, andererseits sollten sie keine zu schwierigen Aufgaben bearbeiten, weil auch das Zeitvergeuden sein kann. Es kann sogar sein, dass Lernende durch zu schwierige Aufgaben falsche Hypothesen über einen Lerngegenstand entwickeln und der Lernprozess dadurch gehemmt wird.

Bei summativen Prüfungen bekommen die Studierenden eine Rückmeldung darüber, ob eine Antwort richtig oder falsch war und warum eine bestimmte Punktzahl vergeben wurde. Bei formativen Prüfungen muss aus der Rückmeldung außerdem hervorgehen, *warum* die Antwort richtig oder falsch war. Bei einer falschen Antwort kann man sein Wissen mit Hilfe dieser Rückmeldung erweitern. Aber auch bei einer richtigen Antwort ist eine Erklärung hilfreich. Es

könnte nämlich sein, dass einzelne Lernende die Antwort einfach geraten haben, oder sie zwar wussten, aber die falsche Überlegung dahinter stand.

Beim formativen Testen ist es – anders als beim summativen Testen – besonders wichtig, dass die Rückmeldung zeitnah erfolgt. Hier haben digitale Tests mit automatischer Korrektur erhebliche Vorteile gegenüber Aufgabenblättern. Man kann sie nämlich so gestalten, dass die Rückmeldung sofort nach Abgabe einer einzelnen Antwort sichtbar wird. Zu diesem Zeitpunkt können sich die Lernenden noch an ihre Überlegungen erinnern und diese durch die Rückmeldung bestätigen oder revidieren. Das neu erworbene Wissen kann bei der nächsten Frage sofort verwendet werden.

Nicht zuletzt sollte am Ende eines formativen Tests immer klar sein, was der nächste Schritt im Lernprozess sein sollte. Sollten die Lernenden nun Aufgaben eines höheren Niveaus bearbeiten oder sollten sie noch auf dem gerade absolvierten Niveau bleiben? Am Ende eines jeden formativen Tests sollten deshalb Empfehlungen stehen. Empfehlungen könnten Lektürehinweise enthalten oder Aufgaben, die als nächstes bearbeitet werden sollten.

Da *time on task* einer der zuverlässigsten Prädiktoren für den Lernerfolg ist (Carol, 1963), sollte formatives Testen Lernende dazu motivieren, sich möglichst viel mit dem Lerngegenstand zu befassen. Anders als beim summativen Testen sollte das Wiederholen von Aufgaben beim formativen Testen deshalb nicht nur möglich sein, sondern auch honoriert werden. Bei analytischen Lerngegenständen, wo es also nicht darauf ankommt, sich ausschließlich Fakten zu merken, ist das Wiederholen von Aufgaben natürlich dann besonders sinnvoll, wenn bei der Wiederholung andere Aufgaben mit ähnlichem Schwierigkeitsgrad angeboten werden.

Die Ausführungen lassen ahnen, dass formatives Prüfen ungleich aufwändiger vorzubereiten ist als summatives Prüfen. Die Aufgaben müssen mit Feedback versehen werden, es muss mehrere Leistungsniveaus geben und idealerweise sollte es mehr Aufgaben geben als die Lernenden unbedingt absolvieren müssen, sodass Tests gewinnbringend wiederholt werden können. Man muss jedoch berücksichtigen, dass der Aufwand sich vor allem bei großen Kohorten von Lernenden nach einiger Zeit amortisiert, weil Korrekturarbeiten – zumindest bei automatischer Korrektur durch die Lernplattform – gänzlich entfallen. Man sollte auch andere positive Effekte des formativen Prüfens bedenken, etwa dass man sehr heterogene Gruppen mit passgenauen Übungen versorgen kann. Verschiedene Studien zeigen, dass formatives Prüfen den Lernerfolg erhöht (u.a. z.B. McDaniel et al., 2007; Kibble, 2007). Auch die folgende Untersuchung zeigt, dass formatives Testen einen entscheidenden Einfluss auf den Lernerfolg hat.

2. Beschreibung des Moduls

Im Folgenden wird ein Modul beschrieben, welches anhand der im vorangegangenen Abschnitt geschilderten Überlegungen zum formativen Testen mittels der Lernplattform Moodle gestaltet wurde. Das Modul wird in der linguistischen Ausbildung von Studierenden des Grundschullehramts an der Humboldt-Universität zu Berlin genutzt. Um zu ermitteln, welche Komponenten dieses Moduls wie stark zum Lernerfolg beitragen, wurde das Modul evaluiert.

2.1. Untersuchte Personengruppe

Untersucht wurden 396 Studierende der Humboldt-Universität zu Berlin mit dem Berufsziel Grundschullehrer(in). Die Studierendengruppe ist in vielerlei Hinsicht heterogen. Die meisten sind zwischen 18 und 45 Jahre alt und bringen sehr unterschiedliche Vorkenntnisse mit. Einige Studierende haben nur recht rudimentäre Kenntnisse zur Grammatik aus der Schule. Andere Studierende haben – meist im Ausland – bereits eine Ausbildung als Deutschlehrer/in absolviert und auch schon in diesem Beruf gearbeitet. Diesen Studierenden sind die Inhalte des Moduls teilweise bekannt. Um Studierende weder zu über- noch zu unterfordern, wurde das Modul sehr flexibel gestaltet, sodass die Studierenden abhängig von ihren Vorkenntnissen genau auf dem Niveau lernen können, wo sie den größten Lernzuwachs haben.

2.2. Materialien

Die untersuchten Studierenden belegten im Untersuchungszeitraum einen einführenden Linguistikkurs. Vorgesehen ist dieser Kurs im 2. Semester, er wird aber von vielen Studierenden bereits im 1. Semester absolviert. Die Daten wurden in einer Zeit erhoben, in welcher aufgrund der Corona-Pandemie kein Präsenz-Unterricht möglich war. Der Online-Unterricht bestand aus einem Seminar, welches über die Konferenzsoftware Zoom durchgeführt wurde. Außerdem wurde über die Plattform Moodle ein Lern- und Übungsmodul mit formativen Tests zur Verfügung gestellt.

Die Bewertung des Kurses gliedert sich in einen formativen Teil, den man durch Erreichen einer bestimmten Punktzahl bestehen muss, und einen summativen Teil, den man ebenfalls bestehen muss, bei dem aber auch eine Bewertung in Form einer Note erfolgt. Der formative Teil wird im Folgenden beschrieben. Der summative Teil besteht aus einer Abschlussprüfung am Ende des Semesters.

Die wöchentlichen Lernangebote des formativen Teils kann man in drei Bausteine einteilen: Vorbereitung, Seminar und Nachbereitung. Bis auf die zweite Phase bearbeiten die Studierenden die Materialien eigenständig außerhalb des Seminars. Am Ende des Semesters gibt es eine weitere Phase, die Klausurvorbereitung. Abbildung 1 visualisiert den Aufbau des Moduls.

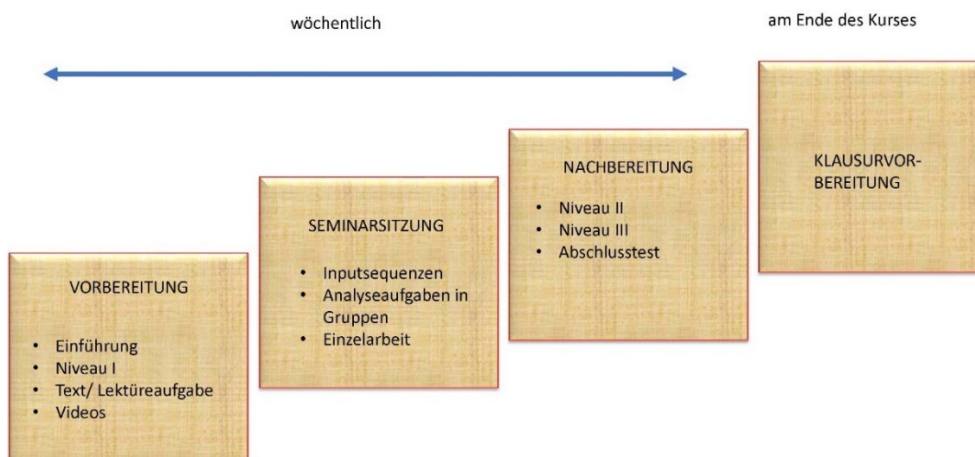


Abb. 1: Aufbau des Moduls mit seinen formativen Elementen

Für fast alle Aktivitäten im formativen Teil gibt es Punkte. Die Studierenden müssen bis zum Ende des Semesters die Hälfte der erreichbaren Punkte erwerben, ansonsten haben sie den formativen Teil des Kurses nicht bestanden. Im Folgenden werden alle Komponenten des formativen Moduls beschrieben.

In der *Vorbereitungsphase* (Abb. 1, Kasten ganz links) haben die Studierenden folgende Lernmöglichkeiten: eine Einführungsaufgabe, die Absolvierung eines formativen Tests (*Niveau I*), einen Lehrbuchtext mit Fragen dazu und Lernvideos. In der Einführungsaufgabe sollen die Studierenden für das Thema interessiert werden und dessen Relevanz für ihren Beruf erkennen. Dazu wird den Studierenden eine möglichst praxisrelevante Frage gestellt, z.B. zur Schreibung eines Wortes. Da es hier vor allem darum geht, ein Bewusstsein für die Problematik zu entwickeln, wird von den Studierenden eine Freitextantwort erwartet. Sobald die Studierenden eine Antwort eingeben, bekommen sie die Punkte für diese Aufgabe. Die Freitextaufgabe wird nicht korrigiert (eine Korrektur wäre nur manuell möglich), aber die Studierenden bekommen eine Rückmeldung in Form eines Lösungsvorschlags, der das thematisierte Phänomen kurz erklärt.

Der Kern der Vorbereitungsphase ist die Aufgabe *Niveau I*. Diese Aufgabe ist technisch betrachtet ein Test, der fast ausschließlich Fragen mit Antwortauswahl enthält (Ankreuzaufgaben, Zuordnungen, Auswahlfelder) und automatisch korrigiert wird. Der Test besteht aus einzelnen Teilaufgaben. In jeder Teilaufgabe wird zunächst ein kurzer, etwa halbseitiger schriftlicher Input gegeben. Anschließend sollen die Studierenden eine Frage beantworten, die sich direkt auf den Input bezieht. Typischerweise wird ein Phänomen an einem sprachlichen Beispiel erläutert und die Studierenden sollen dieses Wissen auf ein anderes sprachliches Beispiel anwenden.

Unmittelbar nachdem die Studierenden eine Antwort eingegeben haben, bekommen sie eine Rückmeldung, aus der hervorgeht, ob ihre Antwort richtig war, was die richtige Antwort gewesen wäre, und warum die eingegebene Antwort richtig oder falsch ist. Die Studierenden sollen mit Hilfe dieser Informationen ihr Wissen erweitern und können es bei der nächsten Frage sofort anwenden. Am Ende des Tests *Niveau I* bekommen die Studierenden einen Hinweis, welche Aktivität ihnen als nächstes empfohlen wird. Wenn sie weniger als 60% der Punkte erreicht haben, lautet die Empfehlung, die Aufgabe *Niveau I* zu wiederholen oder die Videos zu schauen und den Test dann zu wiederholen. Wenn die Studierenden mehr als 60% erreicht haben, wird empfohlen, entweder den Lehrbuchtext zu lesen oder die Videos zu schauen oder in die Nachbereitungsphase zu wechseln und mit der Aufgabe *Niveau II* fortzufahren.

Die dritte Lernaktivität der Vorbereitungsphase (Lehrbuchtext/Lektüreaufgabe) besteht darin, einen Lehrbuchtext zum jeweiligen Thema zu lesen und dazu Fragen zu beantworten. Die Fragen sind typischerweise wieder Fragen, bei denen eine Antwort ausgewählt werden muss. Die Antworten können entweder direkt aus dem Lehrbuchtext entnommen werden oder es handelt sich um einfache Schlussfolgerungen. Die Fragen sollen den Studierenden helfen, festzustellen, ob sie den Text ausreichend gut verstehen. Außerdem sollen die Fragen (und die Punkte, die man hier sammeln kann, wenn die Antwort richtig ist) die Studierenden motivieren, den Text überhaupt zu lesen. Die Lektüreaufgabe gehört zu den wenigen Tests, die nicht wiederholt werden können, denn die Studierenden sollen nicht motiviert werden, die Antworten beim ersten Versuch abzuschreiben und dann einfach einzugeben, sondern dazu, den Text zu lesen und zu verstehen.

Zusätzlich oder als Alternative zu einer der anderen Aktivitäten können die Studierenden sich Lernvideos zu einzelnen häufig als schwierig angesehenen Lerngegenständen ansehen. Bei dieser Aktivität können keine Punkte gesammelt werden.

In der zweiten Phase, der *Seminarsitzung* (Abb. 1, zweiter Kasten von links), bekommen die Studierenden mehrere kurze Inputs mit anschließenden Übungssequenzen. Es wird dabei davon ausgegangen, dass sie sich mit dem Thema der jeweiligen Seminarsitzung schon beschäftigt haben, also die Vorbereitungsphase durchlaufen haben. Im Anschluss werden vor allem Anwendungsaufgaben gelöst, etwa Schreibungen erklärt, Silbenstrukturen und Morphemstrukturen analysiert und ähnliches. Außerdem erhalten die Studierenden immer die Möglichkeit, Fragen zu stellen, die sich in der Vorbereitungsphase ergeben haben. In dieser Phase können keine Punkte gesammelt werden.

Die *Nachbereitungsphase* (Abb. 1, dritter Kasten von links) besteht aus drei unterschiedlichen Lernaktivitäten. Bei der Lernaktivität *Niveau II* sollen die Studierenden das anwenden, was sie im Seminar oder in der Vorbereitungsphase gelernt haben. Die Fragen im Niveau II sind vergleichbar mit denen aus dem Niveau I, es werden aber keine Erklärungen vorab mehr gegeben. Wie im Niveau I erhalten die Studierenden bei richtiger Lösung einen Hinweis dazu, warum diese Lösung richtig ist und bei falscher Lösung einen Hinweis, warum diese Lösung falsch ist. Die Niveau II-Aufgabe kann (und soll) so lange wiederholt werden, bis die Studierenden die Thematik beherrschen. Deshalb ist die Aufgabe technisch so umgesetzt, dass mehr Fragen existieren, als die Studierenden zur Bearbeitung bekommen. Bei jedem Testdurchlauf werden die Fragen zufällig aus einem Aufgabenpool gezogen, sodass bei Wiederholung der Aufgabe meist andere Fragen zu bearbeiten sind. Am Ende der Aufgabe bekommen die Studierenden einen Hinweis, welche Lernaktivität sie als nächstes angehen sollten. Wenn die Studierenden mehr als 60% richtige Antworten gegeben haben, wird ihnen empfohlen, die Aufgabe Niveau III oder den Abschlusstest zu bearbeiten. Bei weniger als 60% richtigen Antworten wird ihnen empfohlen, den Text zu lesen und anschließend die Aufgabe vom Niveau II zu wiederholen.

Bei der Aufgabe *Niveau III* soll das Fachvokabular gefestigt werden. Die Studierenden sehen mehrere kurze Definitionen und sollen dazu den Fachbegriff eingeben. Zum Beispiel sehen sie als Definition: „Laut, der durch einen totalen oralen Verschluss bei gehobenem Velum produziert wird“ und sollen als Antwort „Plosiv“ eingeben. Dadurch, dass die Menge an richtigen Antworten hier definierbar ist (neben *Plosiv* wird auch *Plosive* und *Verschlusslaut/e* akzeptiert), ist eine automatische Korrektur möglich, obwohl es sich nicht um eine Antwort-Auswahl-Aufgabe handelt. Bei der Aufgabe ist ein Zeitlimit von 100 Sekunden eingestellt. Die Aufgabe ist deshalb nur lösbar, wenn man die Begriffe tatsächlich kennt, denn erstens muss man sie eintippen (und nicht nur erkennen, wie bei Ankreuz- oder Zuordnungsaufgaben) und zweitens hat man zeitlich keine Gelegenheit, die Begriffe nachzuschlagen. Auch diese Aufgabe kann und soll wiederholt werden. Die Fragen werden zufällig aus einem Fragenpool gezogen, sodass man zu einem großen Teil andere Fragen bekommt, wenn man wiederholt.

Die Nachbereitungsphase schließt mit einem kleinen Test ab. Die Fragen drehen sich um die wichtigsten Punkte des Themas. Dieser Test kann nicht wiederholt werden. Am Ende des Tests bekommen die Studierenden eine Auswertung, wie gut sie mit ihrem Wissen in der Klausur

abgeschnitten hätten. Der Abschlusstest kann als summatives Element im formativen Teil des Kurses angesehen werden.

Die Studierenden sind relativ frei in der Auswahl der Lernaktivitäten. Es gibt lediglich folgende Einschränkungen:

- Die Einführungsaufgabe muss zuerst bearbeitet werden.
- Die Aufgaben Niveau II und III können nur bearbeitet werden, wenn das jeweils niedrigere Niveau mit 60% bestanden ist.
- Der Abschlusstest kann nur absolviert werden, wenn mindestens das Niveau II absolviert wurde.

Es gibt keine Teilnahmepflicht für das Seminar. Keine der Aufgaben im formativen Modul ist verbindlich. Den Studierenden wird durch dieses Vorgehen die Verantwortung für ihren Lernprozess übertragen (Petko 2010: 11). Sie müssen selbst feststellen, welche Lernaktivitäten ihnen den größten Lernfortschritt bringen und sich dann auf genau diese Aktivitäten konzentrieren.

Das Angebot ist bewusst redundant gestaltet. Beispielsweise ist das Niveau I inhaltlich betrachtet eine komprimierte Version des Lehrbuchttextes. Die Videos greifen einzelne Aspekte auf, die aber auch im Lehrbuch enthalten sind und im Niveau I vorkommen. Auch die Inputsequenzen im Seminar sind inhaltlich eine Zusammenfassung des Lehrbuchttextes. Die Studierenden erhalten also viele Möglichkeiten, sich die Lerngegenstände auf unterschiedliche Weise anzueignen oder das Gelernte zu wiederholen, sodass der Lerneffekt nachhaltiger ist.

Gegen Ende des Semesters haben die Studierenden die Gelegenheit, sich in einen zweiten Moodlekurs einzutragen. Dieser Kurs enthält dieselben Aufgaben wie der Kurs, in welchen die Studierenden während des Semesters eingetragen sind, zusätzlich aber noch eine Beispielklausur. Die Studierenden können hier weiter für die Abschlussprüfung üben, selbst wenn die wöchentlichen Fristen für die Aufgaben im formativen Teil schon abgelaufen sind.

2.3. Ergebnisse der Evaluation

Ziel der Evaluation des formativen Moduls war, herauszufinden, welche Lernaktivitäten von den Studierenden ausgewählt wurden und welche Aktivitäten zum größten Lernerfolg führen. Der Lernerfolg wurde dabei als Punktzahl in der (summativen) Abschlussprüfung gemessen. Im Anschluss an die Evaluation soll das Modul verbessert werden, beispielsweise indem die Studierenden auf besonders erfolgversprechende Lernaktivitäten hingewiesen werden oder diese durch die Vergabe einer höheren Punktzahl attraktiver gemacht werden. Aufgaben, die nicht zur Verbesserung der Leistung beitragen, sollen entfernt werden.

Zunächst wurde erfasst, inwiefern die Aktivitäten überhaupt genutzt werden. Abb. 2–4 zeigen die Anzahl an Versuchen pro Person für die Aufgaben vom Niveau I, II und III. In jeder Abbildung ist auf der horizontalen Achse das wöchentliche Thema angegeben, zu dem der Test angeboten wurde (Themen 1-11). Auf der vertikalen Achse ist die Anzahl an Versuchen pro Person angegeben. Wenn alle Studierenden eine Aufgabe genau einmal absolviert haben, würde das als Strich bei 1 auf

der vertikalen Achse erkennbar sein. Das Niveau I (Abb. 2) des ersten Themas wurde von den meisten Studierenden einmal, von einer relativ großen Gruppe aber auch zweimal absolviert. Einzelne Studierende absolvierten es sogar häufiger, eine Person z.B. fünfmal. Insgesamt kann man feststellen, dass die meisten Studierenden den Test für jedes Thema einmal bearbeitet haben. Nur das letzte Thema haben viele Studierende nicht bearbeitet, was vermutlich daran lag, dass sie die Mindestpunktzahl für den formativen Teil zu diesem Zeitpunkt schon erreicht hatten.

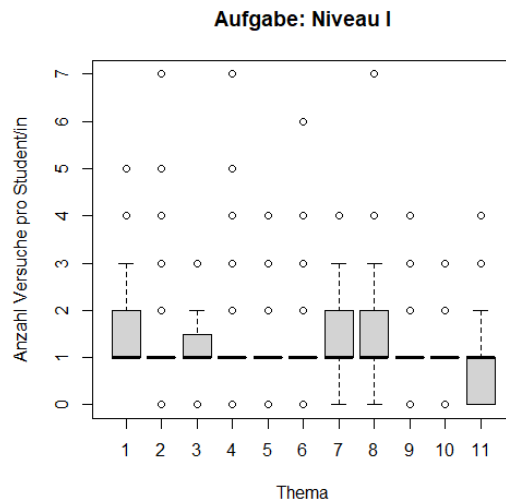


Abb. 2: Anzahl an absolvierten Niveau I-Aufgaben pro Student/-in

In Abb. 3 sind die Ergebnisse für das Niveau II dargestellt. Bei diesem Niveau wurden die Aufgaben zufällig aus einem Aufgabenpool gezogen, sodass Wiederholen sinnvoll war, weil die Studierenden immer wieder andere Aufgaben bekamen. Die horizontale Achse ist genauso skaliert wie bei Abb. 2, die vertikale Achse ist anders skaliert. Es gibt zwar auch bei diesem Niveau zahlreiche Studierende, die das Niveau II nur einmal bearbeitet haben. Die Anzahl an Studierenden, die es mehrfach bearbeitet haben, ist aber deutlich höher als beim Niveau I, die Balken gehen meist bis zur 2. Das bedeutet, dass die Studierenden diese Aufgaben öfter zum wiederholten Üben genutzt haben. Wie auch schon bei Niveau I zu sehen, haben viele das letzte Thema nicht mehr bearbeitet.

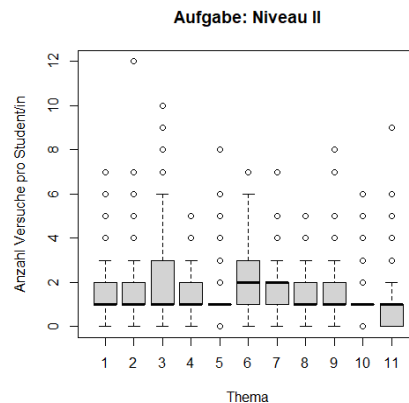


Abb. 3: Anzahl an absolvierten Niveau II-Aufgaben pro Student/-in.

Das Niveau III (Abb. 4) wurde noch häufiger genutzt. Viele Studierende absolvierten diese Aufgabe 2–4mal, einige auch deutlich häufiger (bis zu 22mal).

Abb. 5 zeigt die Resultate für die Leseaufgabe. Diese Aufgabe konnte nur einmal absolviert werden. Gezeigt wird deshalb nicht die Anzahl an Versuchen pro Person, sondern der Prozentsatz an Studierenden, die diese Aufgabe absolviert haben. Die Texte beinhalten jeweils Informationen zu mehreren Themen, weshalb es nicht zu jedem Thema eine Textaufgabe gab. Die Themen 1–3 beschäftigten sich mit der Phonetik, die Themen 4–5 mit der Phonologie, 6–7 mit der Graphematik und 8–11 mit der Morphologie. Abb. 5 zeigt, dass der Lehrbuchtext beim ersten Thema von nahezu allen Studierenden gelesen wurde. Danach sinkt die Anzahl an Studierenden, die den Text lesen und die Fragen beantworten. Die letzte Textaufgabe hat nur noch etwa ein Drittel der Studierenden bearbeitet.

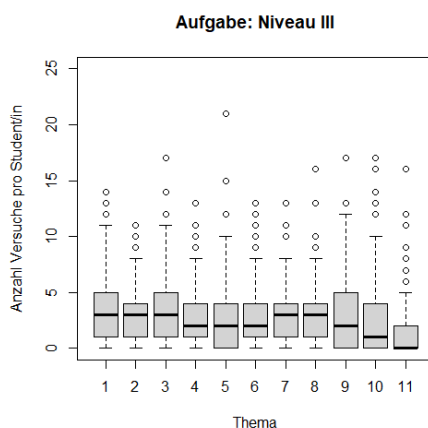


Abb. 4: Anzahl an absolvierten Niveau III-Aufgaben pro Student/-in.

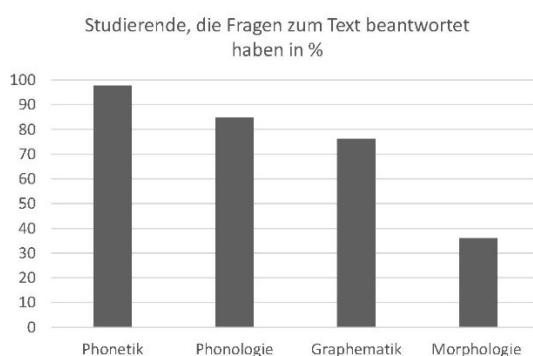


Abb. 5: Prozentsatz an Studierenden, die die Fragen zum Text bearbeitet haben

Abb. 6 zeigt die Ergebnisse für den Abschlusstest. Auch der Abschlusstest konnte nur einmal absolviert werden. Die Entwicklung ist hier ähnlich wie beim Lehrbuchtext. Zu Beginn des Semesters absolvieren knapp 80% diesen Test. Den letzten Test absolviert nicht einmal mehr ein Drittel der Studierenden.

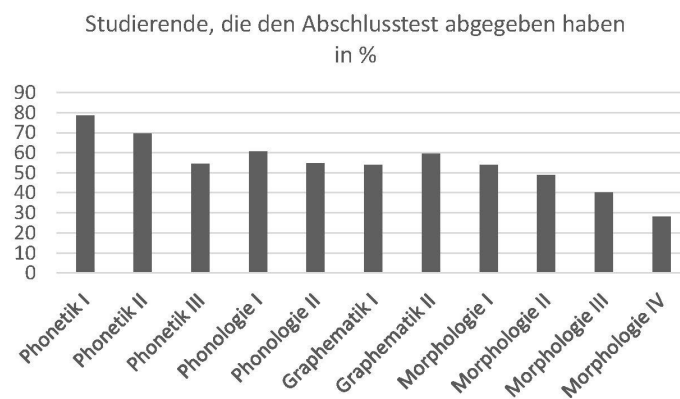


Abb. 6: Prozentsatz an Studierenden, die den Abschlusstest absolviert haben.

Bei der Auswertung der Anwesenheit im Zoom-Seminar wurde festgestellt, dass etwa 10% das Seminar nie besucht haben. Etwa 20% waren nur einmal da. Die anderen verteilten sich etwa gleichmäßig auf die Anzahl der Sitzungen, nur etwa 5% waren immer da.

Die Videos waren offenbar sehr beliebt. Nur 5 Personen haben keins der Videos angeklickt. Ein Viertel der Studierenden hat alle Videos angeklickt.

Mehr als die Hälfte der Studierenden hat den Klausurvorbereitungskurs überhaupt nicht genutzt. Die Studierenden, die ihn genutzt haben, haben sehr unterschiedlich damit gearbeitet. Die meisten haben bestimmte Themen wiederholt, nur sehr wenige (etwa 2%) haben sämtliche Aufgaben wiederholt.

Die Auswertungen zeigen, dass die Studierenden der Empfehlung gefolgt sind, unter den Aktivitäten eine Auswahl zu treffen. Vor allem am Anfang des Semesters probierten sie mehrere Lernaktivitäten aus. Nach kurzer Zeit gingen sie dazu über, nicht mehr alle Aktivitäten zu bearbeiten, sondern vor allem die Aufgaben vom Niveau I, II und III. Die Lernaktivität, die am wenigsten beliebt war, ist die Aufgabe zum Lehrbuchtext.

Der Kurs endet mit einer Abschlussprüfung (summativer Test). Die Fragen in der Abschlussprüfung sind vom Schwierigkeitsgrad her so wie die Aufgaben vom Niveau II. Die Kompetenzen, die geprüft werden sollen, sind vor allem analytischer Natur: Die Studierenden sollen sprachliche Beispiele analysieren. Das heißt, dass keine Aufgabe vom Niveau II direkt in die Abschlussprüfung übernommen wird. Im Unterschied zum formativen Teil des Kurses ist die Abschlussprüfung eine Aufsichtsarbeit. Genauso wie die formativen Tests wird auch die Abschlussprüfung in Moodle geschrieben.

Um zu ermitteln, welche Lernaktivität am stärksten mit der Punktzahl in der Abschlussprüfung zusammenhängt, wurden folgende Daten verwendet:

- die (von den jeweiligen Studierenden erreichte) Punktzahl in allen Aufgaben zu den Lehrbuchtexten
- die Punktzahl in den Niveau I-Aufgaben
- die Punktzahl in den Niveau II-Aufgaben
- die Punktzahl in den Niveau III-Aufgaben

- die Punktzahl in den Abschlusstests
- die Anzahl an Seminarteilnahmen
- die Gesamtpunktzahl im formativen Teil und im Klausurvorbereitungskurs
- die Anzahl an angeklickten Videos.

Mit diesen Variablen wurde ein lineares Modell berechnet mit der in der Abschlussprüfung erreichten Punktzahl als abhängige Variable. Dabei wurde festgestellt, dass die Gesamtpunktzahl und die Teilnahmefrequenz an den Zoom-Seminaren die einzigen Parameter waren, die einen signifikanten Einfluss auf die Punktzahl hatten, die in der Abschlussprüfung erreicht wurde. Der Zusammenhang mit der Gesamtpunktzahl war dabei etwas stärker als der mit der Teilnahmefrequenz. Zusammen erklärten die beiden Variablen knapp 30% der Varianz innerhalb der Daten.

Aus diesem Ergebnis können folgende Schlussfolgerungen gezogen werden. Erstens, die Gesamtpunktzahl ist zwar der wichtigste Einflussfaktor, das Seminar ist für das Lernen aber ebenfalls entscheidend. Dieses Ergebnis spricht dafür, die Studierenden zur Teilnahme am Seminar anzuhalten. Zweitens gab es keine Aufgabentypen, die das Prüfungsergebnis besonders beeinflusst haben. Der Zusammenhang ist viel einfacher: Wer viele Aufgaben absolviert und dabei viele Punkte erreicht, ist auch gut in der Abschlussprüfung.

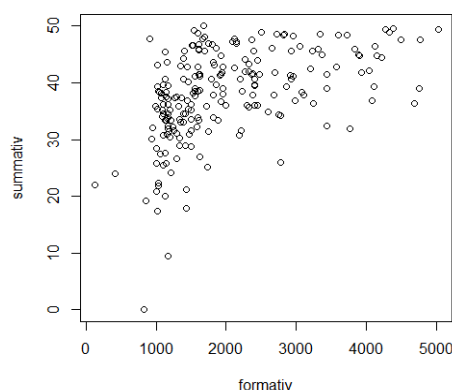


Abb. 7: Zusammenhang zwischen erreichter Gesamtpunktzahl im Kurs (formativ) und Punktzahl in der Abschlussprüfung (summativ).

In Abb. 7 ist dieser Zusammenhang dargestellt. Auf der horizontalen Achse ist die Gesamtpunktzahl angegeben (addierte Punkte im regulären Moodlekurs und im Klausurvorbereitungskurs). Auf der vertikalen Achse ist die Punktzahl in der Abschlussprüfung angegeben (maximal konnten 50 Punkte erreicht werden). Da die Studierenden minimal 1000 Punkte erreichen mussten, um den formativen Teil zu bestehen, haben viele etwas mehr als 1000 Punkte gesammelt. Nur wenige liegen darunter, und diese Studierenden haben in der Abschlussprüfung oft nur wenige Punkte erreicht.

Bei den leistungsstärkeren Studierenden gibt es große Unterschiede bezüglich der Lernaktivitäten. Einige Studierende haben kaum mehr als 1000 Punkte erreicht und schneiden in der Abschlussprüfung sehr gut ab. Andere haben viel mehr geübt und z.B. 4000–5000 Punkte

gesammelt und erreichen trotzdem nur 80% der Punkte in der Klausur. Um zu ermitteln, woran das liegen könnte, wurden weitere, teilweise qualitative, Analysen durchgeführt.

Zunächst wurden die Studierenden betrachtet, die im summativen Test mehr als 40 Punkte erreicht haben, also ein gutes oder sehr gutes Ergebnis. Bei diesen Studierenden wurde festgestellt, dass sie einen Schwerpunkt gelegt haben: Entweder waren sie häufig im Seminar, dann haben sie meist weniger Punkte im formativen Teil erreicht (Punkte oben links in Abb. 7) oder sie waren kaum im Seminar, dann haben sie mehr Punkte im formativen Teil erreicht (oben rechts in Abb. 7). Leistungsschwache Studierende (weniger als 30 Punkte im summativen Teil) waren generell selten im Seminar und haben weniger Punkte erreicht. Leistungsstarke Studierende scheinen also in der Lage zu sein, ihren Lernfortschritt einschätzen zu können. Sie können in der Folge eine sinnvolle Auswahl ihrer Lernaktivitäten vornehmen. Leistungsschwachen Studierenden scheint das nicht zu gelingen.

Als nächstes wurde untersucht, inwiefern die Lernvideos die Unterschiede bei den leistungsstarken Studierenden erklären können. Eine Hypothese war, dass das Lernen durch bloße Rezeption der Videos vielleicht weniger effektiv ist. Die Auswertung der Nutzung der Lernvideos zeigt jedoch nur folgenden Zusammenhang: Diese Videos wurden fast ausschließlich von leistungsstarken Studierenden angeklickt (aber nicht von allen). Einen statistischen Zusammenhang mit der erreichten Punktzahl gibt es innerhalb der leistungsstarken Studierenden nicht. Die leistungsschwachen Studierenden hingegen haben sehr wenige Videos angeklickt. Auch hier gilt also: Wer sich wenig mit den Themen des Seminars beschäftigt, erreicht kein gutes Ergebnis in der Abschlussklausur.

Als dritter weiterführender Aspekt wurde untersucht, ob es sinnvoll ist, Aufgaben zu wiederholen. Der Gedanke dahinter war, dass die Wiederholung von Aufgaben vielleicht gar keinen Lernzuwachs bringt. Es wurden Korrelationen zwischen der Anzahl an abgegebenen Aufgaben eines Niveaus und der Punktzahl in der Abschlussprüfung berechnet. Während bei den Niveau I und II-Aufgaben kein Zusammenhang feststellbar war, war die Korrelation für die Niveau III-Aufgaben überraschenderweise negativ: Je öfter die Studierenden die Aufgabe wiederholt haben, umso schlechter haben sie in der Klausur abgeschnitten. Eine qualitative Untersuchung der abgegebenen Aufgaben zeigte, dass die Studierenden eine starke Tendenz hatten, diese Aufgaben so lange zu bearbeiten, bis sie die volle Punktzahl erreicht haben. Leistungsstärkere Studierende haben dazu weniger Versuche gebraucht als leistungsschwächere. Das begründet die negative Korrelation. Diese Tendenz konnte auch bei den Niveau I und II-Aufgaben beobachtet werden, sie erreichte aber keine statistische Signifikanz.

3. Schlussfolgerungen

Die Ergebnisse der Evaluation des formativen Moduls sprechen sehr für ein *blended learning*-Format in der Hochschullehre, also eine Kombination aus Seminar und digitalen Elementen. Studierende sollten dazu motiviert werden, das Seminar zu besuchen und sich außerhalb des Seminars noch so viel wie möglich mit den Themen des Seminars zu beschäftigen.

Es konnte nicht festgestellt werden, dass einzelne Aufgabentypen überhaupt nicht zum Lernerfolg beigetragen haben. Alle Aufgaben werden von Studierenden genutzt. Am häufigsten

wurden die Aufgaben vom Typ Niveau I, II und III genutzt. Die Studierenden haben eine Tendenz, solange an den Aufgaben zu arbeiten, bis sie ein bestimmtes Niveau erreicht haben. Das bedeutet, dass leistungsschwächere Studierende tendenziell öfter wiederholen als leistungsstärkere. Die Lernvideos wurden vorrangig von leistungsstarken Studierenden genutzt. Auch das Seminar wurde von vielen Studierenden besucht. Am wenigsten wurden die Aufgaben zum Lehrbuchtext genutzt.

Leistungsstarke Studierende profitieren offenbar stark davon, wenn die Anforderungen des Moduls relativ frei gestaltet sind und sie sich Aktivitäten aussuchen können. Leistungsschwächere Studierende würden wahrscheinlich von weniger Freiheiten und einer stärkeren Anleitung profitieren, als sie sie im hier vorgestellten Modul vorgefunden haben. Aufgaben, bei denen man keine Punkte sammeln konnte (z.B. die Videos), haben sie gemieden.

Durch die Wahlmöglichkeiten beim formativen Modul konnten die Bedürfnisse von unterschiedlichen Lerntypen und Ausgangsniveaus besser bedient werden. Das ist besonders bei der hier untersuchten sehr heterogenen Studierendengruppe wichtig. Leistungsschwächere Studierende konnten motiviert werden, Aufgaben zu wiederholen. Einige wenige Studierende kamen sehr gut mit dem Lehrbuchtext zurecht und konnten dabei so viel Informationen entnehmen, dass sie nicht mehr ins Seminar zu kommen brauchten und trotzdem erfolgreich in der Abschlussprüfung waren. Für andere Studierende war das Seminar unverzichtbar.

Man kann beobachten, dass die Studierenden zu einem gewissen Grad durch die Punktvergabe beeinflusst werden: Viele Studierende absolvieren Aufgaben so oft, bis sie die volle Punktzahl erreicht haben. Einige bearbeiten keine Aufgaben mehr, sobald sie die Grenze von 1000 Punkten erreicht haben. Ob eine Aufgabe bearbeitet wird, hängt also nicht ausschließlich davon ab, ob sie von den Studierenden als lernwirksam betrachtet wird, sondern davon, ob sie viele Punkte bringt. Das ist auf der einen Seite bedauerlich, weil es wünschenswert wäre, dass die Studierenden sich mehr Gedanken darüber machen, ob eine Aufgabe gewinnbringend ist. Auf der anderen Seite eröffnet es aber auch die Möglichkeit, die Wahl der Aufgaben stärker zu steuern. Das erscheint vor allem bei den Literaturaufgaben sinnvoll. Lektürekompetenz, auch wenn sie nicht Teil der summativen Prüfung war, ist für das Studium sehr wichtig. In den kommenden Semestern sollen die Lektüretests deshalb durch eine höhere Punktzahl attraktiver gemacht werden. Außerdem erscheint es sinnvoll, die zu erreichende Gesamtpunktzahl etwas zu erhöhen, damit mehr Studierende auch das letzte Thema bearbeiten.

Literaturverzeichnis

- Caroll, John B. (1963): A model of school learning. In: Teachers College Record (64), 723-733.
- Kibble, Jonathan (2007): Use of unsupervised online quizzes as formative assessment in a medical physiology course: effects of incentives on student participation and performance. In: Advances in Physiological Education (31), 253-260.
- McDaniel, Marc A./Janis L. Amderson/Mary H. Derbish/Nova Morrisette (2007): Testing the testing effect in the classroom. In: European Journal of Cognitive Psychology. Vol.19/No. 4-5, 494-513.
- Petko, Dominik (2010): Lernplattformen, E-Learning und Blended Learning in Schulen. In: Petko, Dominik (Hg.): Lernplattformen in Schulen. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 9-27.