

## Teststrategien im virtuellen Deutschunterricht Testing strategies in online German courses

Sebastian Heiduschke<sup>a</sup>

---

### Abstract:

Empirische Studien belegen, dass in virtuellen Kursen häufiger unlautere Mittel eingesetzt werden als im Präsenzunterricht. Dieser Artikel zeigt anhand zweier Fallbeispiele aus dem universitären virtuellen DaF-Unterricht verschiedene Möglichkeiten, die Integrität eines Kurses ohne eine Vielzahl an intrusiven technischen Mitteln zu gewährleisten. Während eine logische Reihung von formativen und summativen Tests den direkten Vergleich studentischer Leistungen erlaubt, kann durch die Kombination asynchroner und synchroner Unterrichtselemente eine Datenbank erstellt werden, welche als Grundlage zur Bewertung von Tests dient.

### Keywords:

Digitalisierung, Evaluation, online lernen, Teststrategien, virtuelle Lehre

---

### Resumen:

Empirical studies show a higher frequency of cheating in virtual courses than in face-to-face classes. This article uses two case studies of university-level virtual German language classes to show less-intrusive, non-technical methods that still ensure the course integrity. Whereas sequencing formative and summative tests appropriately allows the direct comparison of student performance, a combination of asynchronous and synchronous assignments can be used to create a database for the evaluation of test results.

### Palabras Clave:

Digitalization, evaluation, online learning, testing strategies, virtual teaching

---

### VORÜBERLEGUNGEN

Spätestens als sich viele Universitäten in den USA ab dem Frühjahr bedingt durch Covid-19 gezwungen sahen, den Präsenzunterricht innerhalb von kürzester Zeit in den digitalen Bereich zu verlagern, wurde vielen Deutschlehrer(\*)innen der Nutzen einer gutdurchdachten Digitalstrategie für Deutsch als Fremdsprache bewusst. Programme, die neben Präsenzunterricht bereits hybride Formate und Onlineunterricht anboten, konnten sofort auf diese Unterrichtsformate zurückgreifen und Kurse umwandeln. Student(\*)innen, die bereits Erfahrungen mit dem virtuellen Fremdsprachenunterricht hatten, konnten auf diese zurückgreifen und mussten nicht erst technologisch geschult werden. Hatte man bis vor einem Jahrzehnt universitäre digitale DaF-Kurse häufig als undurchführbar titulierte und mit deren Entwicklung zumindest auf universitärer Ebene aufgeschoben, sahen sich nun Lehrkräfte vor vollendete Tatsachen gestellt, besonders als klar wurde, dass es sich bei der Pandemie nicht um eine kurzfristige Erscheinung handelte,

nach der man zum vorherigen Status Quo des Präsenzunterrichts zurückkehren würde. [1] Videokonferenzen ersetzten asynchrone Übergangslösungen wie Blogs, Arbeitsblätter oder statische Seiten auf diversen Lernplattformen, das Home-Office den Kursraum.

Neben negativen Effekten gab es jedoch auch eine positive Seite. War die weltweite DaF-Gemeinde schon immer für ihre Kreativität und Enthusiasmus bekannt, gesellte sich nun Solidarität hinzu. Man suchte die globale Vernetzung, wandte sich über Facebookgruppen und WhatsApp Chaträume mehr als zuvor an Kolleg(\*)innen, teilte Materialsammlungen und offerierte Erfolgsgeschichten, bot Videoseminare und virtuelle DaF-Fortbildungen an. [2] Der Umstieg auf das virtuelle Lernen, egal ob synchron, asynchron oder als Hybridform, wurde durch das didaktisierte Material zumindest erleichtert, so dass man sich schneller wieder den Lernenden zuwenden konnte. Jedoch war auch dort oft Hilfestellung vonnöten, da die Student(\*)innen weniger „Digital Wisdom“—um hier Marc Prenskys eigene Differenzierung der ursprünglichen Dichotomie von „Digital Natives“ und „Digital Immigrants“

<sup>a</sup> Sebastian Heiduschke, Oregon State University, Email: [sebastian.heiduschke@oregonstate.edu](mailto:sebastian.heiduschke@oregonstate.edu)

zu zitieren—besaßen, um selbständig und eigenverantwortlich die Lernmaterialien auf der Lernplattform durchzuarbeiten. [3] Die wichtigsten Fragen zum digitalen DaF-Unterricht scheinen geklärt, ganz im Gegensatz zum Testen, das in nahezu jedem Webinar zum virtuellen Lehren und Lernen angesprochen wird.[4] Konkret geht es vielen Lehrkräften gewöhnlich um Fragen in Bezug auf unerlaubte Hilfsmittel, da die Befürchtung besteht, dass sich Kursteilnehmer(\*innen) in virtuellen Kursen mangels direkter Aufsicht häufiger unredlicher Mittel bedienen, und dass dadurch die Integrität der Kurse sowie die Erfüllung der Lernziele gefährdet seien. Es gibt in der Tat empirische Hinweise auf eben diese Gefahren; eine vergleichende Studie ergab, dass Student(\*innen) in virtuellen Kursen zwölfmal öfter schummelten als in Präsenzkursen.[5] Gerade in Kursen, die in ein virtuelles Format transformiert wurden, ohne sie der anderen Lehrumgebung anzupassen, beobachtete man oft, dass bis zu 25% der Getesteten unerlaubte Hilfsmittel verwendeten. [6]

Trotz oder vielleicht gerade wegen der ernüchternden Realität ist zu überlegen, welche Schritte vonnöten sind, um das Testen in virtuellen Kursen so zu gestalten, dass die Prüfungsergebnisse genauso verlässlich sind wie im Präsenzunterricht. Da in keinem Unterrichtsmodus die Tests hieb- und stichfest sein werden, egal welche Mittel wir anwenden, ist unser Ziel die Reduzierung des Mogelns auf ein möglichst geringes Maß, damit die Schummelquoten in unseren virtuellen Kursen denen des Präsenzunterrichts gleichen. Manche Kolleg(\*innen) fordern zwar verpflichtend die Einführung von Prüfungen in virtuellen Kursen unter Aufsicht [7], übersehen dabei aber, dass diese Kurse unter anderen Lehr- und Lernbedingungen stattfinden als unsere Präsenzkurse.[8] Sobald wir den ersten großen Schritt getan haben, indem wir unsere Kursinhalte den Begebenheiten des virtuellen Unterrichts angepasst haben—einige Stichworte hier sind Flipped Classroom, Zoom-Bombing, Ausgewogenheit von asynchronen und synchronen Elementen—können wir uns dem Testen zuwenden.[9]

## LERNZIELANPASSUNG

Auch hier gibt es Vorgehensweisen, welche die Integrität eines virtuellen Kurses verbessern können. Im Folgenden stellt dieser Artikel anhand zweier Fallstudien vor, wie die Germanistische Fakultät unserer Universität ihre Tests und Evaluationen der Student(\*innen) im virtuellen DaF-Unterricht gestaltet hat. Hierbei handelt es sich um eine Kombination aus formativen und summativen Evaluationsmethoden, die wir im letzten Jahrzehnt für unseren virtuellen DaF-Unterricht im Rahmen unserer universitären Onlinelehre entwickelt haben. Es ist anzumerken, dass wir gerade mit der summativ-formativen Kombination den größten Erfolg erzielt haben und sich die Täuschungsrate auf etwa zwei Prozent eingependelt hat. [10]

Bevor wir uns den Testmethoden zuwenden, soll zunächst angemerkt werden, dass es sich bei der ersten Fallstudie um einen Sprachkurs auf dem B1 Niveau des Gemeinsamen Europäischen Referenzrahmens handelt, während mein zweites Beispiel aus einer Serie von Kursen im Rahmen unseres Programms „Technisches Deutsch“ besteht, in denen Student(\*innen) auf den Niveaustufen A1 bis B2 unterrichtet werden. Diese Sequenz ist aber inhaltlich auf allen Stufen gleich konzeptualisiert, so dass wir sie in einer Fallstudie zusammenfassen.

Der B1-Kurs trägt bei uns den Namen „Drittes Jahr Deutsch“. Unsere Student(\*innen) durchlaufen die gesamte Sequenz vom ersten bis zum vierten Jahr für den Bachelorabschluss auf dem Niveau B2/C1 oder bis zum dritten Jahr für das Nebenfach Deutsch auf der Stufe B1/B2, haben allerdings je nach Vorerfahrung die Möglichkeit, sofort im zweiten (A2) oder dritten Jahr (B1) zu beginnen. Wir unterrichten im Trimesterzyklus und erwarten einen wöchentlichen Arbeitsaufwand von ca. 12 Stunden, also insgesamt 120 Stunden pro Trimester. In unseren virtuellen Kursen treffen wir uns viermal pro Trimester in Kleingruppen von bis zu fünf Student(\*innen) für eine synchrone mündliche Übung, während die Hybridversion des B1 Kurses zweimal pro Woche eine synchrone einstündige Sitzung im gesamten Klassenverband vorsieht, was bedeutet, dass unter Umständen bis zu 25 Student(\*innen) gleichzeitig im virtuellen Zoom-Klassenzimmer des Hybridkurses sitzen. Die B1 Kurse verwenden als Lehrwerk Netzwerk B1 von Klett (Kursbuch und Arbeitsbuch), welches sowohl in gedruckter Form als auch als App verfügbar ist, die auch ohne Datenverbrauch oder ständige Internetverbindung genutzt werden kann. Wir benutzen Canvas als universitäre Lernplattform und veröffentlichen dort den Wochenplan und die Zugangsdaten für unsere Videokonferenzen.

Die Serie „Technisches Deutsch“ ist eine Kollaboration zwischen den Abteilungen Werkstoffwissenschaften und Germanistik. Student(\*innen) haben die Möglichkeit, einen dualen Bachelorabschluss in Mechanical Engineering von unserer Universität und in Werkstoffwissenschaften einer deutschen Partneruniversität zu erhalten. Dazu studieren sie drei Jahre an unserer Universität und bereiten sich studienbegleitend auf den Auslandsaufenthalt vor, indem sie Technisches Deutsch belegen. Hier haben wir jedoch nur drei Stunden wöchentlich zur Verfügung, wovon nur eine Stunde für ein synchrones Treffen genutzt werden kann. Jede Niveaustufe muss also mit relativ wenig Kontaktstunden bewältigt werden. Auf der anderen Seite belaufen sich die Einschreibezahlen durch die Spezialisierung der Student(\*innen) auf weniger als ein Dutzend pro Niveaustufe, was intensives und zügiges Arbeiten in den Zoom-Treffen erlaubt. Die Kurse verwenden alle drei Komponenten (Lernen, Machen, Zeigen) des Lehrwerks Impuls Deutsch 1 bzw. Impuls Deutsch 2 ausschließlich als E-Buch, das über eine App oder das Internet zugänglich ist.

## EVALUATIONSMETHODEN

Die unterschiedlichen Rahmenbedingungen erforderten eine Anpassung unserer Teststrategien an die jeweilige Lernumgebung. Unsere Teststrategien erlauben uns, Daten zu erheben um den Wissenstand unserer Student(\*))innen im Blick zu behalten, um korrektiv eingreifen und individuell fördern zu können. Dabei orientieren wir uns wie bereits erwähnt am Europäischen Referenzrahmen.[11]

Da die Lernziele unseres Lehrplans durch die jeweiligen „Kann-Beschreibungen“ auf den diversen Niveaustufen vorgezeichnet sind, passten wir unsere Kursziele zunächst der Globalskala des jeweiligen Sprachniveaus an, formulieren anschließend die modularen Lernziele ebenfalls als „Kann-Beschreibungen“ und unterteilen diese in die zwei Kategorien Hören und Sprechen sowie Lesen und Schreiben.[12] Während der B1 Kurs Lernziele wie „Kann über Musik und Komponisten sprechen und schreiben“ hat, finden sich im Kurs „Technisches Deutsch“ hingegen Ziele wie „Kann über den eigenen ökologischen Fußabdruck sprechen und schreiben“. Hier wird deutlich, dass sich die Inhalte unterscheiden mögen, die prinzipiellen methodologischen Überlegungen zum Testen aber weiterhin gelten. Ausgehend davon liegen also Handlungsorientierung und die Kompetenzorientierung unserer Teststruktur als inhaltliche und formale Gesichtspunkte zu Grunde. Unsere Zielsetzung ist, unseren Student(\*))innen zum einen zu ermöglichen, Sprachprüfungen des Goethe-Instituts oder von telc, bzw. das ÖSD zu bestehen, und zum anderen, sich selbstständig im deutschsprachigen Sprachraum zu bewegen und dort eine berufliche Tätigkeit ausüben zu können. Als Konsequenz ergab sich für uns die Erstellung von Rubriken, um den Lernerfolg der Student(\*))innen zu evaluieren. Dabei folgten wir den Bewertungskriterien der bereits genannten Examen und trennten die einzelnen Prüfungen in Teilprüfungen, die unabhängig voneinander abgelegt und bestanden werden können.

## FORMATIVES TESTEN

Beginnend mit dem B1-Kurs soll in diesem Aufsatz zunächst auf die einzelnen Testkomponenten eingegangen werden. In B1 führen wir in jedem Modul über den Zeitraum von zweieinhalb Wochen das neue Material schrittweise im Rhythmus von zwei bis drei Tagen mit Hilfe der Lernplattform Canvas ein. Dort haben die Student(\*))innen Zugang zu interaktiven Lehrvideos, die begleitend zu unserem Buch grammatikalische Konzepte einführen und sofort einen unbewerteten Verständnistest durchführen. Je nach Konzept handelt es sich hier entweder um Lückentexte, Multiple Choice oder die Ordnung syntaktischer Elemente anhand der neu eingeführten Regel. Das Video kann nach eigenem Belieben abgespielt werden und das Wissen risikolos geprüft werden. Gewöhnlich kombinieren wir diese Videos mit

Aufgaben aus dem Arbeitsbuch, die selbständig erledigt und mit Hilfe des Lösungsschlüssels von den Student(\*))innen eigenständig überprüft werden. Auch hier vergeben wir keine Punkte, stellen aber durch verbale Instruktionen klar, dass diese Aufgaben notwendige Schritte sind und der Lernfortschritt am Ende des Moduls durch summative Tests geprüft wird, welche in Format und Inhalt den Übungen ähneln. Selbständiges Arbeiten und Eigenmotivation sind für den Lernerfolg in virtuellen Kursen notwendig. Oft genügt der Hinweis auf die Parallelen zwischen den formativen Übungen und den summativen Tests als Stimulus für extrinsisch Motivierte.

Jeder der Schritte innerhalb eines Modul ist noch einmal unterteilt in einen Einzelauftrag und in eine Aufgabe, die mit einer Partner(\*))in oder in einer Kleingruppe erledigt werden. Die Aufgaben entnehmen wir entweder dem Lehrbuch und dem Arbeitsbuch, um so zur Festigung des Wortschatzes und der Anwendung neuer Phrasen und Grammatik beizutragen. Hier bearbeiten die Student(\*))innen die gesamte Bandbreite an Aufgaben, seien es geschlossene Formate wie Lückentexte und Vokabeln via Quizlet, oder aber offene Formate wie das Schreiben von Mails, Geschichten, oder Blogs. Wir mischen die Aufgaben, so dass die Student(\*))innen voneinander lernen, wenn sie Rückmeldungen zu Fehlern in den Aufgaben der anderen geben, oder wenn die Gruppe zusammen an einer Schreibaufgabe arbeitet und bereits präventiv Fehler beim gemeinsamen Verfassen vermeidet.

Gruppen- und Partneraufgaben müssen stets auf Video aufgezeichnet und auf die Lernplattform hochgeladen werden, offene Einzelaufgaben hingegen werden schriftlich in einem kursinternen Diskussionsraum auf Canvas so gepostet, dass die Aufgabe und die Korrekturen der Lehrkraft für alle Student(\*))innen sichtbar ist, die Note aber nur der oder dem jeweiligen Student(\*))in zugänglich ist. Wir setzen diese Schritte als formative Evaluation ein und bewerten mit Hilfe einer dreistufigen Rubrik, welche die Leistung in die Kategorien „komplett und sorgfältig erledigt“, „teilweise erledigt“ und „mangelhaft erledigt“ klassifiziert. Obwohl die Aufnahmen es erlauben würden, die Entwicklung der Sprachmittlungskompetenz zu überprüfen, da wir im Laufe eines Kurses mehr als zwanzig Sprachaufnahmen pro Student(\*))in zur Verfügung haben, beschränkt sich unsere Bewertung nur auf die drei Stufen. Wir bauen damit Angst vor Fehlern ab und belohnen Risiko.

Einzelaufgaben bewerten wir variierter, zum Beispiel wenn Student(\*))innen ihren Lernfortschritt bei der Wortschatzerweiterung dokumentieren, indem sie am Ende einer Woche das Vokabular mit Hilfe eines von ihnen selbst erstellten Quizlets nach unseren Vorgaben mit 20 Fragen testen und einen Screenshot mit dem Ergebnis hochladen. Dieses Ergebnis tragen wir direkt in das Notenbuch ein. Da dieses Quiz beliebig oft wiederholt werden darf, erzielen die meisten Student(\*))innen hohe Punktzahlen. Der Effekt dieses

Evaluationselements ist durchaus vorteilhaft: unsere Student(\*))innen verfügen über ein großes Vokabular und können es in den Diskussionen selbständig anwenden, was wir in den summativen Tests sehen, in denen wir das aktive Beherrschen eines extensiven Vokabulars in mehreren Teilbereichen bewerten. Da die Quizlets nur acht Prozent der Endnote ausmachen und wir zudem den Erfolg unserer Methode beobachten können, glauben wir, dass hier keine Verfälschung vorliegt.

Offene Formate werden ausführlich mit einem zweistufigen Ansatz bewertet. Zum einen verwenden wir angelehnt an das Bewertungsschema des Goethe Instituts die vier Rubriken „Erfüllung der Aufgabenstellung“, „Wortschatz“, „Strukturen“ und „Kohärenz“ und evaluieren auf diese Weise die Leistung mit einem numerischen Wert. Wir können dadurch früh Probleme identifizieren und der/dem Student(\*))in kommentarlos vermitteln, in welchen Bereichen Übungsbedarf besteht. In diesen offenen Formaten etablieren die Student(\*))innen gleichzeitig eine Ausgangsbasis und liefern uns einen Textkorpus, den wir später zum Vergleich mit den summativen Leistungen heranziehen können. Hier ist anzumerken, dass fast alle unserer Lehrkräfte zertifizierte Tester(\*))innen für die Goethe Sprachtests sind, wir also die nötige Ausbildung besitzen, um zusätzlich zu den Inhalten auch eine authentische Testatmosphäre, z.B. mittels Körpersprache oder verbaler Rückmeldung, zu schaffen.

Zusätzlich zu dieser Note erhalten die Student(\*))innen ein personalisiertes Video, in dem wir mittels Screenshare das Geschriebene besprechen, Probleme zeigen, Lösungsansätze bieten und, wenn notwendig, Hilfestellung leisten. In diesen Videos bietet sich die Gelegenheit, darauf hinzuweisen, wenn wir sehen, dass Student(\*))innen unerlaubterweise mit Übersetzungsprogrammen gearbeitet haben. Wir bieten in diesem Fall an, die Aufgabe eigenständig zu wiederholen. Damit schaffen wir eine Vertrauensbasis, wir helfen statt zu bestrafen, wir zeigen aber auch, dass wir Plagiate identifizieren können. Wir verwenden Screenshare, um den Text mit Hilfe des Cursors mitzuverfolgen, aber auch um weniger bedrohend zu wirken. Unser Unterrichtsalltag unterstreicht den Erfolg dieses Vorgehens. Wir sehen an den Übungsaufgaben, dass die Student(\*))innen bereits vor der ersten summativen Evaluation die formelhaften Wendungen und Strukturen von E-Mail und Blogbeitrag sicher beherrschen, und sich vollkommen den Rubriken Struktur und Kohärenz widmen können.

Obwohl unsere Universität vorgibt, Kurse so weit wie möglich asynchron zu unterrichten, bieten wir pro Modul ein synchrones Treffen als Standard an. Die Student(\*))innen tragen sich vorab für ein einstündiges Zoomtreffen in Kleingruppen von maximal fünf Teilnehmer(\*))innen ein. Wir bieten diese Treffen gewöhnlich über zwei Tage zu verschiedenen Tageszeiten an, weil unsere Student(\*))innen die Möglichkeit des virtuellen Studiums weltweit nützen. Es ist nicht ungewöhnlich, dass in einem Kurs mehrere Kontinente

und Altersgruppen vertreten sind, und dass Schüler(\*))innen, Berufstätige, Zeitsoldat(\*))innen und Rentner(\*))innen einander in spannende Konversationen verwickeln. Für den Fall, dass kein Termin passt, erlauben wir, die Aufgaben asynchron zu erledigen und auf die Lernplattform zu laden. In der Praxis ist das aber nur selten der Fall; die Kursteilnehmer(\*))innen erwähnen in ihren Kursbewertungen, dass die synchronen Treffen ein Highlight der Kurse darstellen und den größten Lernerfolg bringen. Diese Treffen ähneln im Aufbau dem Präsenzunterricht, nur dass sie per Zoom durchgeführt werden. Im Idealfall erlaubt die Begrenzung auf eine kleine Gruppe eine durchschnittliche Redezeit von etwa neun Minuten pro Person, also deutlich mehr und in konzentrierterer Form als es bei uns im Präsenzunterricht möglich ist, wenn deutlich mehr Personen im Kursraum sitzen. Die Treffen werden aufgezeichnet und stehen der Lehrkraft im Anschluss zur Verfügung. Wir können uns also auf den Ablauf konzentrieren, gezielt verbessern, und müssen erst in der Nachbereitung mit Hilfe der Aufnahme evaluieren und können anschließend unsere detaillierten Erkenntnisse per Videobotschaft mit der/dem Student(\*))in teilen. In dieser Hinsicht ist es uns tatsächlich möglich, die Ergebnisse der formativen Tests so einzusetzen, dass Verbesserungen vor der summativen Evaluierung am Modulende möglich sind. Zudem gilt auch hier, dass die Aufnahmen uns kontrastives Vergleichsmaterial liefern, mit dem wir eventuellen Betrug bei den formativen Tests erkennen können.

## **SUMMATIVES TESTEN**

Ein Block aus sechs summativen Tests zum Hörverstehen, Schreiben, Leseverständnis, Sprechen, Grammatik und Vokabular schließt jedes Modul ab, wobei wir Teilkompetenzen isoliert prüfen. Zuvor führen die Student(\*))innen eine Selbstevaluation durch, in der sie für sich selbst bewerten, wie sicher jedes Lernziel erreicht wurde, die von der/dem Student(\*))in angekreuzt werden. Wir arbeiten hier mit vier Emojis, von lachend bis unzufrieden, um die „Kann“-Aussagen zu kommentieren. Die Evaluation wird nicht mit der Lehrkraft geteilt; stattdessen geben wir begleitende Hinweise, in welchen Kursschritten das jeweilige Lernziel behandelt wurde, um eine Wiederholung des Materials zu erleichtern. Dann öffnen wir am Freitag ein dreitägiges Testfenster, innerhalb dessen die Student(\*))innen die Tests in beliebiger Reihenfolge bearbeiten können. Die Ergebnisse werden erst am Montag zugänglich gemacht, wenn alle die Prüfungen abgelegt haben.

Ein Pilotprojekt, in dem wir diese Teilprüfungen zu einem langen Test zusammenfassten, erwies sich nach einer informellen Befragung der Student(\*))innen eines gesamten Jahrgangs als leistungshemmend. Von den 30 Student(\*))innen sprachen sich fast alle gegen eine lange Prüfung aus.

Manche bemängelten, dass bei technischen Schwierigkeiten der gesamte Test wiederholt werden musste, da eine sofortige

Problembhebung oft nicht möglich war, wenn die Student(\*))innen die Aufgaben zu einer unkonventionellen Uhrzeit angingen. Für die Lehrkraft war es zudem nicht möglich, mit Gewissheit zu sagen, inwieweit die Aufgabe bereits eingesehen worden war, und bis zu welchem Grad der Test kompromittiert war. Weitere negative Faktoren die genannt wurden, waren Stress, wenn äußere Einflüsse zum Beispiel im häuslichen Bereich die Student(\*))innen zwangen, den Test zu unterbrechen und viel Zeit verloren, die nicht mehr aufgeholt werden konnte. Für viele war es schlichtweg unmöglich, einen durchgehenden Block von bis zu zwei Stunden pro Modul zu reservieren, währenddessen sie ungestört arbeiten konnten. Trotz mehrfacher Änderungen ergab sich nach Abschluss des zweijährigen Pilotversuchs, dass diskrete Tests die besten Prüfbedingungen für den virtuellen Kurs darstellten. Zum einen finden die Student(\*))innen nun bessere Prüfungsbedingungen vor, da manche Tests in fünf Minuten abgeschlossen werden können, während wir für längere Tests wie Aufsätze, Mails oder Blogposts zwar ein Zeitlimit vorgeben, aber ein Puffer für unvorhergesehene Zwischenfälle auf Seiten der Student(\*))innen einkalkulieren. Zum anderen spiegelt die flexible Zeiteinteilung bei den Tests besser die asynchrone Lernumgebung der Kurse wider, die ihrerseits durch die modulare Struktur mit vielen diskreten Arbeitsanweisungen den Student(\*))innen Spielraum zur Eigenständigkeit zugesteht.

Die sechs Tests werden auf der Lernplattform Canvas eingestellt und per Kalenderfunktion automatisch geöffnet beziehungsweise geschlossen. Für jeden Test setzen wir eine Zeitbeschränkung, die dem Umfang, Format und Schwierigkeitsgrad der Aufgabe angepasst ist. Ein Multiple-Choice Vokabeltest mit fünf Drop-Down-Listen muss nicht länger als fünf Minuten dauern, während wir für das Schreiben einer Mail auf dem B1-Niveau eine Gesamtdauer von 45 Minuten ansetzen. Wir verwenden die gesamte Bandbreite an geschlossenen Formaten wie zum Beispiel Multiple-Choice und Lückentexte, offenen Formaten wie die bereits genannte Mail oder einen Blogeintrag, aber auch dichotome Formate (meist richtig/falsch), diese jedoch generell in Kombination mit anderen geschlossenen Formaten, um den Nachteil der dichotomen Formate auszubalancieren, da selbst bei Nichtwissen mit 50%iger Wahrscheinlichkeit die korrekte Antwort erraten werden kann. Jede korrekte Antwort ist einen Punkt wert, was zur Folge hat, dass ein Vokabeltest gleich viel zählen kann wie eine offene Frage, die frei beantwortet werden muss. Interessant ist allerdings, dass dies auf dem B1 Niveau von unseren Student(\*))innen akzeptiert, ja sogar honoriert wird, weil ein schlechter Test in einem Teilbereich nicht unbedingt eine schlechte Kursnote bedeuten muss, und mit einem guten Test in einem „einfachen“ Bereich problemlos ausgeglichen werden kann. Im Gegenzug scheinen uns die Student(\*))innen mit Aufrichtigkeit zu danken, da wir

selbst bei Grammatiktests oder Aufsätzen niedrige Betrugsquoten feststellen.

Wie stellen wir aber den korrekten Ablauf der Tests sicher, wenn diese virtuell und ohne Aufsicht durchgeführt werden? Die Antwort mag zunächst überraschen, aber wir verzichten weitestgehend auf technische Hilfen, die durchaus verfügbar wären, und verlassen uns stattdessen auf einfachste Verfahren, die es uns und unseren Student(\*))innen ermöglichen, die summative Evaluierung so stressfrei und reibungslos wie möglich durchzuführen. Die Dauer ist knapp kalkuliert, aber dennoch so, dass selbst eine langsame Internetverbindung keine Schwierigkeiten darstellt. Besonders für Übungen in geschlossenen Formaten bieten wir eine Aufgabensammlung an, aus der eine bestimmte Anzahl vom Computer zufällig ausgewählten Übungen bearbeitet werden muss, so dass jede Person unterschiedliche Aufgaben erhält. Bei Aufgaben mit Auswahlantworten werden die Optionen vom Computer jedes Mal gemischt, so dass die Dropdowns für jede Person in einer anderen Reihenfolge erscheinen. In einer weiteren Variation werden sogar die Aufgaben zusätzlich gemischt. Wir wenden dasselbe Prinzip auch bei offenen Formaten an, oder wir überlassen den Prüflingen die Wahl, indem wir eine Arbeitsanweisung erstellen und eine Auswahl an Themen oder Bildern, über die gesprochen bzw. geschrieben werden soll, offerieren. Als Hilfsmittel für unsere Bewertung stellt unsere Lernplattform ein Quiz-Logbuch zur Verfügung, auf dem wir den Testablauf nachvollziehen können. Wir sehen, wie lange ein\*e Student(\*))in an einer Aufgabe gearbeitet hat, wir werden benachrichtigt, ob verbotenerweise ein anderes Browserfenster oder ein anderes Tab geöffnet wurde, und wir können die Tippgeschwindigkeit und Buchstabenfolge bei offenen Aufgaben sehen, und sogar, ob aus einem externen Programm kopiert und der Text in das Testfeld eingefügt wurde. In diesem Fall bewerten wir die Aufgabe nicht, informieren die Student(\*))in und beginnen ein obligatorisches, von der Universität eingeführtes Disziplinarverfahren, in dem eine Konferenz mit dem Dekanat, der Student(\*))in und der Lehrkraft anberaunt wird. Wir weisen auf dieses Verfahren vor jedem Test hin und erläutern die Funktionsweise des Logbuchs. Zudem haben wir die Daten aus den formativen Tests, die wir zur Testbewertung heranziehen können. Als Fakultät an einer US-amerikanischen Universität unterliegen die Prüfungsdaten in einem Disziplinarfall nicht dem Datenschutz und können als Beweismaterial angeführt werden. Die Kombination aus Datensammlung und Prävention hat sich bewährt. Im Kurs B1 mussten in den vergangenen Jahren keine Disziplinarverfahren stattfinden.

An dieser Stelle sei trotzdem auf eine Vielzahl an technischen Hilfsmitteln hingewiesen, die wir erwogen haben, aber letztendlich aus unterschiedlichen Gründen nicht zum Einsatz bringen. Unsere Lernplattform ist standardmäßig mit der Plagiatsoftware Turnitin ausgestattet, die aber nur bei der Bewertung von Mails hilfreich wäre. Dort verlassen wir uns jedoch auf unsere eigenen Daten, da wir relativ schnell die

verschiedenen Schreibstile und typischen Fehler von Student(\*))innen erkennen. Eine automatische Prüfungsaufsicht durch Proctorio wäre umsonst verfügbar, erfordert jedoch, dass ein Test auf einem Computer mit einem bestimmten Browser (Chrome oder Edge) und einer vorinstallierten Browsererweiterung abgelegt werden muss, da Webcam und Mikrophon von dieser Erweiterung während eines Tests gesteuert werden. Ähnliche Verfahren sind unter den Namen „Lockdown Browser“ und „Respondus Monitor“ bekannt. Weil die Fehleranfälligkeit dieser Programme hoch ist, und Bewegungen im Hintergrund durch Kinder oder Haustiere als Probleme markiert werden, ist ein hoher Arbeitsaufwand im Anschluss an die Prüfung erforderlich. Zudem verlassen sich viele unsere Student(\*))innen auf ihre mobilen Endgeräte und erwarten, dass sie Tests auch unter widrigen Bedingungen ablegen können. Deshalb scheidet für uns diese Überwachungsmöglichkeiten aus, ebenso wie eine Beschränkung der Tests auf unsere Campus IP-Adressen. Diese könnten über unser universitäres Virtual Private Network von jedem studentischen Endgerät angewählt werden, der Ansatz würde aber am Technologieverständnis vieler Student(\*))innen scheitern. Statt eines Einsatzes von computergestützter Überwachungsmethoden bringen wir zumindest bisher erfolgreich eine Kombination formativer und summativer Tests zum Einsatz.

### **FORMATIVES UND SUMMATIVES TESTEN ALS UNTERRICHTSMODELL**

Das Testen in der Serie „Technisches Deutsch“ stellte uns aufgrund der schwierigen Rahmenbedingungen vor besondere Herausforderungen. Ziel der Serie ist es, mit nur neunzig synchronen Kontaktstunden und zusätzlichen 180 Stunden für Hausaufgaben und Evaluierung die kommunikative Kompetenz auf dem Niveau B1 in der Fachsprache Werkstoffwissenschaften zu erlangen. Eine zusätzliche Schwierigkeit ist, dass die Kurse nur in einem Bewertungssystem von „Bestanden/ nicht bestanden“ belegt werden können, die Student(\*))innen also intrinsisch motiviert sein müssen, um die von uns gesetzten Bezugspunkte zu erreichen. Glücklicherweise ist diese Motivation dadurch gegeben, dass der duale Abschluss nur dann erlangt werden kann, wenn der Auslandsstudienteil angetreten und erfolgreich abgeschlossen wird. Da von unserer Seite die Empfehlung zum Studium an der deutschen Universität ausgesprochen wird und wir die Sprachfertigkeiten attestieren, ergeben sich sehr selten größere Probleme, zudem alle Student(\*))innen in Deutschland noch einen vierwöchigen Intensivkurs mit Abschlussprüfung am Goethe-Institut absolvieren. Dennoch war es notwendig, die Teststruktur den äußeren Gegebenheiten anzupassen und jede sich bietende Gelegenheit zur Evaluierung des Lernfortschritts zu nutzen.

Den Spagat zwischen Lehren und Evaluieren lösen wir, indem jede Aktivität zu einem Test wird. Anders als im regulären

Sprachunterricht gibt es keine unbewerteten Aufgaben. Dazu war es erforderlich, die Kursbücher nur als E-Books zu erlauben, denn wir bieten bei einem Großteil der geschlossenen Aufgaben die Möglichkeit, sich die Fehler vom Computer anzeigen zu lassen, und die fehlerhaften Antworten zweimal zu korrigieren. Wir weisen die Student(\*))innen wiederholt darauf hin, dass die Konsultation der Lehrwerke bei allen Tests nicht nur erlaubt, sondern ausdrücklich erwünscht ist, da diese Art der lernprozessbegleitenden Diagnose den Student(\*))innen aus ihrem Hauptfach der Werkstoffwissenschaften nicht unbedingt bekannt ist. Hier setzen wir also die Methoden des Low-Stakes-Testens ein, um als Endresultat eine Note zu erhalten, die mit einem summativen High-Stakes-Test vergleichbar ist. Möglich gemacht wird dieser Prozess durch die Computerunterstützung der Lernplattform Blink Learning, auf der die drei Kursbücher den Student(\*))innen zugänglich gemacht werden. Wir nutzen Canvas in dieser Kursserie, um Termine, Aufgaben und Zugangsinformationen zu den synchronen wöchentlichen Treffen mitzuteilen, mit den Student(\*))innen zu kommunizieren, und die Testergebnisse einzutragen, aber nicht zum Testen.

Das Lehrwerk Impuls Deutsch bietet uns durch seine gut durchdachte Struktur, seine englischsprachigen Grammatikerklärungen, das Flipped-Classroom Modell und eine Dreiteilung in ein Buch zum Wissenserwerb (mit dem Titel Lernen), ein Buch für die Arbeit im Klassenverband (Machen, und ein Buch zur summativen Evaluierung (Zeigen) einen nützlichen Rahmen für die Unterrichtsgestaltung. Uns war wichtig, Student(\*))innen der Naturwissenschaften Regeln und eine logische Struktur zu präsentieren, mit deren Hilfe man sich das notwendige Basiswissen bereits vor den Treffen aneignen kann. Hilfreich ist auch, dass dieses Wissen sofort durch die bereits erwähnten formativen geschlossenen Fragen kontrolliert werden kann und durch den Computer korrigiert wird. Zusätzlich zur Grammatik gibt es kulturelle Informationen und Fragen in offenen Formaten, die zur Reflexion über das Gelesene und Gelernte einladen. Besonders komplexe Ideen dürfen auch auf englisch beantwortet werden, wovon interessanterweise nur Student(\*))innen auf dem A1 oder A2 Niveau Gebrauch machen. Die Student(\*))innen im dritten Jahr antworten fast ausnahmslos auf deutsch, da sie wissen, dass nur der Inhalt, jedoch nicht die Grammatik bewertet wird. Schließlich beinhaltet die Komponente Lernen auch offene Formate, die die Teilkompetenzen oft isolieren, manchmal aber auch integrativ testen. Die Lernplattform bietet hier neben Textfeldern auch eine Vielzahl an Aufnahmemöglichkeiten an, um den mündlichen Ausdruck zu üben. Die Beurteilung dieser Aufnahmen nehmen wir vor den synchronen Treffen vor, um dort direkt auf Probleme aus den formativen Tests eingehen zu können. Leider bietet das Lehrwerk keine Möglichkeit zur direkten Antwort per Korrekturvideo an, und anders als z.B. Moodle bietet die Lernplattform Canvas zumindest im Moment keine einfache

Integration der E-Books, so dass wir etwas umständlicher vorgehen müssen, indem wir unsere Videoantworten zuerst auf den Medienserver unserer Universität hochladen und dann den Link darauf in das Kommentarfeld eintragen, der dann von dort von den Student(\*))innen in einem Browser geöffnet wird. Ein Vorteil ist allerdings, dass wir dadurch Zugriff auf eine Statistik haben und sehen können, wie oft das Video von der/dem jeweiligen Student(\*))in angesehen wurde.

Ein Vorteil der zum Großteil automatisierten formativen Evaluierung ist neben der sofortigen Rückmeldung an die Student(\*))innen eine niedrigere Hemmschwelle bei den Videoaufnahmen. Die Möglichkeit, diese ohne direkte Präsenz der Lehrkraft so oft zu wiederholen, bis sie mit dem Ergebnis zufrieden sind, hat den Effekt, dass in den virtuellen Kursen deutlich mehr geübt wird als im Präsenzunterricht. Wenn anschließend die Korrektur durch die Lehrkraft erfolgt ist und die Student(\*))innen das Korrekturvideo mehrmals vor dem synchronen Treffen ansehen konnten, ist in den wöchentlichen Videokonferenzen deutlich weniger Nervosität unter den Teilnehmer(\*))innen zu spüren, selbst wenn hier im gesamten Klassenverband gearbeitet wird und nicht nur in einer Kleingruppe wie im B1 Kurs. Der virtuelle Kursraum birgt noch einen weiteren Vorteil: wir verlangen nicht zwingend, dass die Kamera eingeschaltet werden muss, so dass die Hemmschwelle besonders bei schüchternen Student(\*))innen noch weiter sinkt. Damit eliminieren wir Scham über das häusliche Lebensumfeld, geben den Teilnehmer(\*))innen mit der Kontrolle über die eigene Kamera auch die psychologisch wichtige Kontrolle über ihre Erscheinung, und schaffen so einen Raum, der eine angenehme und kollegiale Atmosphäre für den Spracherwerb schafft. Weiterhin bewerten wir die synchronen Treffen nicht mit einer prozentualen Skala wie wir es mit den formativen Vorbereitungsaufgaben und den summativen Nachbereitungstests tun, sondern honorieren bereits die Teilnahme und Anstrengung. Auch hier zahlt sich Kulanz aus. Ganz selten erleben wir, dass ein(\*))e Student(\*))in einem virtuellen Treffen unentschuldig fernbleibt, sondern selbst bei unzureichender Vorbereitung teilnimmt, und auf diese Weise zumindest passiv einen Nutzen aus der Konversation zieht.

Wie bereits erwähnt unterscheidet sich die summative Evaluation deutlich von anderen Kursen in unserem Curriculum. Wir verzichten in der Serie „Technisches Deutsch“ aus Zeitgründen auf als solche gekennzeichnete Tests und greifen stattdessen auf die summativen Übungen aus dem dritten Lehrwerk „Zeigen“ zurück. Dabei gehen wir wie in der formativen Evaluation in „Lernen“ vor und tragen den prozentualen Durchschnittswert aus den bearbeiteten Aufgaben in unser Notenbuch. Unsere Kurspläne weisen explizit darauf hin, dass die erzielten Noten aus jedem der drei Teilbereiche ein Drittel der Gesamtnote ausmachen, und dass das dritte Buch dazu verwendet wird, zu bewerten, ob das Material erlernt wurde. Auch hier verzichten wir aber nicht auf

die Selbstevaluation in der Mitte und am Ende des Kurses, wenn die Student(\*))innen eine Checkliste abhaken, auf der die Lernziele der vergangenen Wochen verzeichnet sind. Als Lehrkräfte können wir fehlende Häkchen kommentieren und mit der/dem Student(\*))in Defizite noch einmal ansprechen.

Die größten Nachteile an diesem Verfahren sind die fehlende Möglichkeit, die Dauer der summativen Tests festzulegen, und keine Kontrollmöglichkeit über die verwendeten Hilfsmittel zu haben. Theoretisch ist es möglich, dass Student(\*))innen die Antworten zu den Fragen nachschlagen und dann eintragen. Die App zeigt zwar an, wie lange ein(\*))e Student(\*))in mit der jeweiligen Frage verbracht hat, was uns eventuelle Rückschlüsse erlaubt, aber bei weitem nicht in dem Maße wie es mit unserer Lernplattform möglich ist. Die einfachste Lösung wäre, die summativen Tests auf die Lernplattform zu verlegen. Wir gehen aber auch hier einen anderen Weg. Da der Studiengang aufgrund der hohen Arbeitsbelastung intrinsische Motivation verlangt, und am Ende nicht die Scheine für erfolgreich belegte Deutschkurse zählen, sondern die Zulassung zum Auslandsstudium, belassen wir es bei dem gegenwärtigen System. Im ungünstigsten Fall wandelt sich die summative Evaluierung zu einer formativen, wenn sich Student(\*))innen ihrer Notizen und Lehrbücher bedienen, um die Antworten zu geschlossenen Fragen zu finden. Im Falle von offenen Formaten stellt sich das Problem nicht; hier haben wir, wie im B1 Kurs, unsere Datensammlung und können vergleichen, ob es Abweichungen gibt, die auf externe Hilfsmittel wie Übersetzungsprogramme schließen lassen.

Mit dem weitestgehenden Verzicht auf Peer-Evaluierungen weist die Serie „Technisches Deutsch“ eine weitere Besonderheit auf. Während im B1 Kurs diese Art von Evaluierung durch die wöchentliche Partner- bzw. Gruppenarbeit die zentrale Komponente ist, kommt es aufgrund der niedrigen Anzahl an Kontaktstunden vergleichsweise selten zu einer gegenseitigen Kontrolle durch Kommiliton(\*))innen. In den synchronen Treffen arbeiten wir hauptsächlich mit der gesamten Gruppe; dort interveniert die Lehrkraft bei Fehlern und bittet die Teilnehmer(\*))innen um Korrektur. Ist Partner(\*))innen- oder Gruppenarbeit geplant, teilen wir die Student(\*))innen in virtuelle Gruppenräume auf und zirkulieren dann durch die Gruppen, beantworten Fragen, berichtigen Fehler und evaluieren die Leistungen. Dies geschieht ohne Einfluss auf die Note, sondern dient der formativen Evaluation, wenn wir im Plenum die Fehler der einzelnen Gruppen für alle zusammenfassen und verbessern.

Bisher hat uns das Testen in dieser Serie vor keine größeren Herausforderungen gestellt, obwohl wir wie im B1 Kurs keine Überwachungstechnologie verwenden. Wir verlassen uns auch hier auf den Datenkorpus, den uns die Student(\*))innen durch die formativen Tests an die Hand geben. In dieser Hinsicht bietet uns das Lehrwerk sogar präzisere Daten: der Lernfortschritt in den Kategorien Lesen, Hören, Schreiben, Sprechen, Vokabular, Grammatik, Aussprache und Reflexion wird sofort auf dem Bildschirm als Graph mit einem

prozentualen Wert angezeigt. Das Lehrwerk bietet theoretisch das Erstellen eigener Tests und die Einbettung in die Lernplattform, was wir aus Zeitgründen in diesem Kurs nicht nützen, wir aber bei Verwendung in zukünftigen Kursen durchaus in Betracht ziehen würden. Für die Serie „Technisches Deutsch“ hat sich unser Konzept des Unterrichtens durch formatives und summatives Testen bewährt.

## FAZIT

Kehren wir abschließend zu unseren Vorüberlegungen zur Testgestaltung im virtuellen Unterricht zurück. Ist das Testen in diesen Kursen tatsächlich schwieriger oder wurde es bisher mit den Mitteln angegangen, die aus dem Präsenzunterricht bekannt sind? Als die Online-Lehre noch in den Kinderschuhen steckte, musste man die Methoden an das neue Umfeld anpassen. Erst als der Flipped Classroom die traditionelle Vorlesung ablöste, asynchrone Elemente mit synchronen Treffen kombiniert wurden, und Aufgaben so verfasst wurden, dass sie mit einer Vielzahl moderner Endgeräte, unterschiedlichen Betriebssystemen und variabler Softwareumgebung problemlos gelöst werden konnten, konnte man die Vorteile des virtuellen Klassenzimmers nutzen. Für viele hat dieser Prozess erst begonnen und wird eine Weile in Anspruch nehmen.

Was wäre, wenn man das Testen mit einer offenen Haltung angehe und überlegt, ob man mit ähnlichen Mitteln wie der Unterrichtsstruktur und Kursgestaltung bessere Erfolge erzielen könnte als zu versuchen, die Testumgebung aus dem Präsenzunterricht zu replizieren? Mehr Technologie muss nicht immer die bessere Option sein. Tun wir den Student(\*))innen einen Gefallen und erreichen wir tatsächlich authentischere Ergebnisse, wenn wir intrusive Überwachungselemente wie Kameras und Mikrofone einsetzen, und Fremden einen Einblick in die Privatsphäre unserer Lerner(\*))innen erlauben? Eine mögliche Alternative wäre, die Informationen besser zu verarbeiten, die tagtäglich im Fremdsprachenunterricht gewonnen werden. Besonders im virtuellen Unterricht böte sich an, die Daten aus formativen Aufgaben zu sammeln, zu klassifizieren und dann mit den Ergebnissen aus den summativen Tests abzugleichen. Dabei muss nicht jedes Detail aus der Datenfülle der asynchronen Aufgaben penibel beachtet werden. Oft genügt es, bei offensichtlichen Warnzeichen das erfasste Material heranzuziehen und entsprechende Maßnahmen zu ergreifen. Entsprechende Kursplanung macht es möglich, über Daten aus allen sechs Bereichen Sprechen und Schreiben, Lesen und Hören, Vokabular und Grammatik zu verfügen. Andere Alternativen wären, die Gewichtung zwischen formativen und summativen Tests so zu verschieben, dass der Anreiz zum Betrug verkleinert wird, oder die intrinsische Motivation zu fördern, indem Tests authentischer werden, wie zum Beispiel durch Konversationen mit Muttersprachler(\*))innen auf

Plattformen wie TalkAbroad. Andere Ansätze wären empirische Projekte gefolgt von einem Referat und einer schriftlichen Abschlussarbeit, aber auch Tests, die aktiv von Student(\*))innen mitgestaltet werden oder eine Selbstbewertung beinhalten. Viele Wege führen ans Ziel. Absolvent(\*))innen in unserem virtuellen Bachelorprogramm schließen mit einem im Durchschnitt höheren Sprachniveau ab als ihre Kommiliton(\*))innen, die reinen Präsenzunterricht absolviert haben, erhalten Stellenangebote in ihren Wunschberufen, studieren an unseren Partneruniversitäten in Deutschland und bestehen standardisierte Sprachtests. Innovation zahlt sich aus, besonders im noch jungen Feld des virtuellen Sprachunterrichts.

## REFERENCES

- [1] Ein im Jahr 2017 von der Webseite Inside Higher Ed erstelltes Booklet enthüllt weitreichende Skepsis zum Technologieeinsatz in der Universitätslehre. Das Booklet ist kostenlos zu bestellen unter <https://www.insidehighered.com/booklet/2017-survey-faculty-attitudes-technology>.
- [2] Zur Thematik siehe hier zum Beispiel das IDV-Magazin 98. Dezember 2020. <https://idvnetz.org/wp-content/uploads/2020/12/idv-magazin-dezember-2020.pdf>.
- [3] Prensky, M. (2001a). Digital natives, digital immigrants. *On the Horizon* 9(5), 1-6. Prensky, M. (2001b). Digital natives, digital immigrants, Part II. *On the Horizon* 9(6), 1-6. Prensky, M. (2009). H. sapiens digital: From digital immigrants to digital natives to digital wisdom. *Innovate* Volume 5, Issue 3 Article 1.
- [4] Siehe hier u.a. die Datenbanken des Goethe-Instituts [digu.goethe.de](http://digu.goethe.de) und [digfo.goethe.de](http://digfo.goethe.de), das südostasiatische Projekt EDDU (Erfolgreich digital Deutsch unterrichten), das südosteuropäische Projekt DIP (Digital-Interkulturell-Projektorientiert), die transnational digitale Netzwerkuniversität [www.netzwerkuni.com](http://www.netzwerkuni.com). Arnold P., Kilian, L., Thilloßen, A., Zimmer, G. (2011). Handbuch E-Learning. Lehren und Lernen mit digitalen Medien. Bielefeld: Bertelsmann. Roche, J. (2008). Mediendidaktik. Ismaning: Hueber. Rösler, D. (2010). Die Funktion von Medien im Deutsch als Fremd- und Deutsch als Zweitsprache-Unterricht. In H.J, Krumm. *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch* (1199-1214). Berlin: de Gruyter.
- [5] Lucky, A., Branham, M., Atchison, R. (2019). Collection-based education by distance and face to face: Learning outcomes and academic dishonesty. *Journal of Science Education and Technology*, 28, 414-428.
- [6] Newton D. (2020, August 7). Another problem with shifting education online: cheating. <https://hechingerreport.org/another-problem-with-shifting-education-online-cheating/>.
- [7] Ibid.
- [8] Hilfreich ist hier die Datenbank der Oregon State University Ecampus Abteilung, die kontrastive Studien zur Effektivität

katalogisiert.

<https://ecampus.oregonstate.edu/research/projects/online-learning-efficacy-research/>

- [9] Boettcher, J., Conrad, R. (2021). *The Online Teaching Survival Guide: Simple and Practical Pedagogical Tips*. 3rd edition. San Francisco: Josey-Bass. O'Keefe, L. Rafferty, J. Gunder, A., Vignare, K. (2020, May 18). *Delivering high-quality instruction online in response to COVID-19: Faculty playbook*. Every Learner Everywhere. <http://www.everylearnereverywhere.org/resources>
- [10] Dieser Prozentsatz basiert auf unseren Einschreibezahlen und den Disziplinarverfahren, die von unserer Fakultät eingeleitet werden.
- [11] Hier sei besonders auf das von ALTE (2011) produzierte Pamphlet *Manual for Language Test Development and Examining* hingewiesen.
- [12] <https://www.europaeischer-referenzrahmen.de/sprachniveau.php>

## **BIBLIOGRAPHY**

Anderson, L. W., & Krathwohl, David R. (2001). *A taxonomy for learning, teaching, and assessing : a revision of Bloom's taxonomy of educational objectives (Complete ed.)*. Longman.

Boettcher, J., Conrad, R. (2021). *The Online Teaching Survival Guide: Simple and Practical Pedagogical Tips*. 3rd edition. San Francisco: Josey-Bass.

Crosslin, M., Austin, A., Mandal, S., Clark, A., Dellinger, J.T., Dombrowsky, T., Sattler, M., & Cavallo, A. (2020, April 13). *Measuring student learning in online courses: assessment strategies that work*. [Powerpoint slides]. <https://utacrtle.org/docs/measuring-student-learning.pdf>

Crosslin, M., Dellinger, J. T., Milikic, N., Jovic, I., & Breuer, K. (2019). Determining learning pathway choices utilizing process mining analysis on clickstream data in a traditional college course. In J. Cunningham, N. Hoover, S. Hsiao, G. Lynch, K. McCarthy, C. Brooks, R. Ferguson, & U. Hoppe (Eds.), *Companion Proceedings of the 9th International Conference on Learning Analytics and Knowledge* (pp. 231-232). Tempe, AZ: Society for Learning Analytics Research. [http://solar.quadsimia.biz/wp-content/uploads/2019/08/LAK19\\_Companion\\_Proceedings.pdf](http://solar.quadsimia.biz/wp-content/uploads/2019/08/LAK19_Companion_Proceedings.pdf)

Newton D. (2020, August 7). *Another problem with shifting education online: cheating*. <https://hechingerreport.org/another-problem-with-shifting-education-online-cheating/>

O'Keefe, L. Rafferty, J. Gunder, A., Vignare, K. (2020, May 18). *Delivering high-quality instruction online in response to COVID-19: Faculty playbook*. Every Learner Everywhere. <http://www.everylearnereverywhere.org/resources>

Panadero E. (2017). *A review of self-regulated learning: Six models and four directions for research*. *Frontiers in Psychology*, 8, 422. <https://doi.org/10.3389/fpsyg.2017.00422>