Formatives Assessment im ETHZ-Basisjahr

ALEXANDER CASPAR (ETH ZÜRICH)

Online-Multiple-Choice-Aufgaben sind ein fester Bestandteil des Curriculums in den mathematischen ETHZ-Grundvorlesungen im Basisjahr natur- und ingenieurwissenschaftlicher Studiengänge. Sie bieten den Studienanfänger/-innen ein wirkungsvolles und effizientes Instrument zum Üben und zur Selbsteinschätzung. Zudem erhalten Dozierende umfassende Daten über Stand und Entwicklung ihrer Studierenden. Wir präsentieren und diskutieren quantitative und qualitative Ergebnisse des Einsatzes in der Mathematikausbildung an der ETH Zürich. Im Vordergrund steht dabei ein Selbsteinschätzungstest über Schulmathematik, welchen seit 2009 jeweils mehr als 50 % ETHZ-Studienanfänger bearbeiten.

1. Einleitung

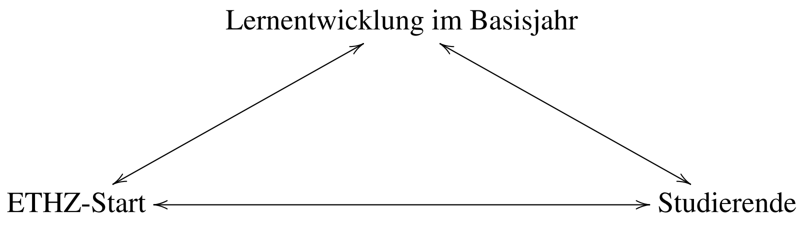
An der Eidgenössischen Technischen Hochschule Zürich (ETHZ) nimmt die Mathematikausbildung eine zentrale Rolle ein: Nahezu alle Studierenden besuchen mathematische Vorlesungen in den ersten Studienjahren. Das Erreichen der Lernziele entscheidet mit über das weitere Studium. Das Departement Mathematik der ETHZ (D-MATH) ist für die Mathematikausbildung verantwortlich. Angesichts der stark gestiegenen Studierendenzahl stand und steht es vor der Herausforderung, die mathematische Ausbildung auf dem gewohnten Niveau sicherzustellen.

In den letzten Jahren konnte das duale System aus Vorlesung und Übungen durch die Neuen Medien um weitere didaktisch-methodische Prinzipien ergänzt werden (vgl. Caspar, Miller (2012)). Die Szenarien mussten dabei die Qualität des Unterrichts für mehr als 10’000 Studierende pro Jahr sichern. Unter Berücksichtigung der unterschiedlichen Voraussetzungen, Lernorte, Spezialisierungen usw. (vgl. Niegemann (2004), S. 71ff.) wurden Online-Multiple-Choice-Aufgaben (MC) als formatives Assessment zur Selbsteinschätzung etabliert (vgl. Kerres (2001), S. 207).

Dazu bedurfte es der

- Entwicklung mathematikspezifischer Standards und Designs von MC-Aufgaben,
- Definition von Umsetzungsszenarien, die sich den Bedürfnissen der Studierenden und Dozierenden anpassen lassen,
- Entwicklung eigener Systeme für individuelle und pragmatische Lösungen (vgl. Kerres, de Witt (2004), S. 77ff.) mit optimaler Passung zwischen Konzept und Nutzungskontext (vgl. Reinmann (2004), S. 111f.).

Nach einer intensiven Aufbau-, Implementierungs- und Etablierungsphase untersuchen wir nun die Wechselwirkungen gegenseitiger Rückmeldungen durch die MC-Aufgaben:

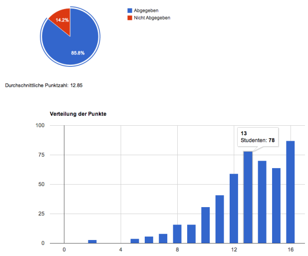


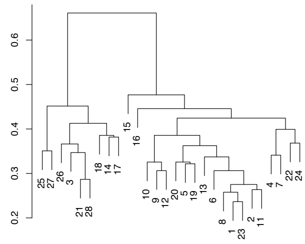
Die Datenerhebung umfasst dabei bisher fünf Studierendengenerationen mit pro Schuljahr

Vorlesungen	Studierende	Übungsserien	MC-Aufgaben
31	$> 10'000$	> 400	$> 1'500$

Schriftenreihe zur Didaktik der Mathematik der Österreichischen Mathematischen Gesellschaft (ÖMG), Heft 48, 2015, S 30 - 45

Mit elementaren Methoden[1] der Statistik/Testtheorie





Punkte- Fehler- Struktur-Verteilung
Erfolgs- Item- Cluster-Analyse

suchen wir nach geeigneten Aufgaben mit Blick auf Motivation, Lernfortschritt und Prüfungserfolg. In einem ersten Schritt analysieren wir hier Fragen eines Selbsteinschätzungsstests über Schulmathematik.

2. Selbsteinschätzungstest Schulwissen

Nach D. P. Ausubel ist “Der wichtigste Faktor, der das Lernen beeinflusst, [ist] das, was der Lernende bereits weiss[2].”

> “If I had to reduce all of educational psychology to just one principle, I would say this: The most important single factor influencing learning is what the learner aleady knows”
> (Ausubel (1968), VI)

In der Mathematik erfährt dieser Grundsatz noch mehr Bedeutung, da neues Wissen grundsätzlich an Vorwissen anknüpft. Solide mathematische Grundkenntnisse der aufeinander aufbauenden Objekte und Methoden entscheiden über den erfolgreichen Studienverlauf. Die Studierenden rekrutieren sich aus divergenten Gymnasien mit unterschiedlichen Stundendotationen (bis zu einem Faktor zwei), Oberstufenjahren, disziplinären Profilen und weiteren Spezialitäten. Daraus resultieren heterogene Eingangsvoraussetzungen.

Wie steht es um das Vorwissen der Studienanfänger an der ETHZ?

Dies können ETHZ-Studienanfänger/innen seit dem Herbstsemester 2009 auf freiwilliger Basis im Rahmen eines Selbsteinschätzungstests mathematisches Schulwissen (S21t) überprüfen.

In einer mathematischen Grundlagenvorlesung am D-MATH erhalten sie in der ersten Semesterwoche eine Einladung zu einem Online-Selbsteinschätzungstest mit 28 (± 1) MC-Fragen. Die Rücklaufquote lag bisher bei 50 – 60% mit einem Maximum in 2013 mit knapp 1′900 Teilnehmenden (63 %).

2.1. Zweck

Primär bietet der S21t die Möglichkeit einer Selbsteinschätzung zum Aufdecken von Lücken und Defiziten, zum Beispiel nach einem Zwischenjahr, Militärbesuch usw. – auch mit Blick auf heterogene CH-Schullandschaft. Selbsteinschätzung des eigenen Leistungsniveaus ist Spitzenreiter in der Hattie-Visible-Learning-Studie (vgl. Hattie (2013), S. 52ff.)

Im Idealfalls motivieren die Rückmeldungen die Teilnehmenden sogar zu einer eigenständigen Repetition. Neben einer ausführlichen Literaturliste gibt es im Moment keine ergänzende Hilfestellung. Ein Brückenkursprojekt wird dies ab Sommer 2016 auffangen.

[1] Neben Erfahrung/Anstrengung

[2] Obwohl der Autor die Allgemeine Hochschulreife in Deutschland erlangt hat, und dieser Artikel in Österreich erscheint, verwenden wir die Schweizer Rechtschreibung. Bestehende Fehler sind davon unabhängig.

2.2. Nebeneffekte

Nach wiederholten erfolgreichen Durchläufen stellten sich auch wichtige Nebeneffekte heraus. Die Ergebnisse und Rückmeldungen der Studierenden erlauben ein Monitoring des Eintrittsniveaus und ermöglichen eine Identifikation von Parametern für den Studienerfolg. Mittlerweile unterstützt der Test jeweils die Prozesse der ETHZ-Aufnahmeprüfung und der Studienberatung.

2.3. S21t-Fragen

Der Test unterliegt keinen wissenschaftlichen Standards der Testkonstruktion. Die Zusammenstellung der Fragen orientiert sich an vorgängigen Tests am D-MATH, Erfahrungen am Lehrstuhl Mathematik und Ausbildung und weiteren Studien (HSGYM[3]).

Die Fragen verteilen sich auf

Algebra, Trigonometrie, Funktionen, Folgen & Reihen, Differential- und Integralrechnung, Analytische Geometrie in Ebene und Raum, Kombinatorik und Wahrscheinlichkeitsrechnung

und sind dabei in einer dreistufigen Taxonomie strukuriert: Der Lösungsweg einer Aufgabe ergibt sich

KOG1 durch Wissen.

KOG2 durch elementare Überlegungen.

KOG3 durch anspruchsvollere Überlegungen.

Die folgende Tabelle zeigt die Einsortierung der Fragen - jeweils als • - in die drei Stufen:

Gebiet	KOG1	KOG2	KOG3	Summe
Algebra	•	• •	•	4
Trigonometrie	•	•	• •	4
Funktionen	•	• •	• •	5
Folgen, Reihen		• •		2
Differentialrechnung	•	•	• • •	5
Integralrechnung	•	•	• •	4
Geometrie Ebene/Raum	•	• •	• •	5
Kombinatorik		•		1
Wahrscheinlichkeitsrechnung		•	•	2
Summe	6	13	13	32

Dabei sind Doppelnennungen möglich. Zum Beispiel gehören Fragen zum Verhalten trigonometrischer Funktionen in die Gebiete Trigonometrie und Funktionen.

2.4. Modus Operandi

Die Studierenden belegen eine mathematische Grundlagenvorlesung in einem von 25 Bsc-Studiengängen der Ingenieur- und Naturwissenschaften[4]. In der ersten Semesterwoche erhalten sie eine Einladung zum Online-Test. Empfohlen wird, die Aufgaben mit Papier und Stift als einzige Hilfsmittel – und Köpfchen – zu bearbeiten und dafür 50 bis 60 Minuten einzuplanen.

Nach der Bearbeitung und Eingabe erhalten sie Rückmeldungen via Lösungen und Einordnung der eigenen Leistung im Vergleich zur Gesamtheit der Teilnehmenden. Eine globale Auswertung und Statistik folgt nach dem Einsendeschluss, welcher zehn Tage später angesetzt ist.

[3] Im HSGYM-Projekt arbeiten mehrere Hundert Zürcher Mittelschullehrpersonen und Dozierende der Zürcher Hochschulen an einer Verbesserung der Schnittstelle Gymnasium – Hochschule, Link http://www.hsgym.ch/startseite.

[4] Hier sind die Studierenden der Mathematik und Physik mit weitergehenden mathematischen Lernzielen ausgeklammert.

2.5. Evaluation

In den Durchgängen 2009, 2010 und 2013 hatten wir durch eine Online-Befragung Nutzen und Einsatz des Tests untersuchen können. Mit einem Rücklauf von > 50% der Teilnehmenden (TN) lagen jeweils aussagekräftige Beobachtungen vor.

Quantiative Rückmeldung der Studierenden

Jeweils ergab sich ein ähnliches Bild.

- Rund 70 – 80% der TN können mit den Fragen ihr Schulwissen einschätzen.
- Die resultierende und eigentlich intendierte Motivation zur Repetition ist unentschieden.
- Knapp ein Viertel benötigt zusätzliche Hilfsmittel (neben Papier und Stift und Köpfchen).
- Durchschnittliche Bearbeitungszeit liegt in der Tat zwischen 50 und 60 Minuten.
- Rund 90% der TN empfehlen den Test anderen Studierenden.

Qualitative Rückmeldung der Studierenden

Neben den quantitativen Rückmeldungen zeigen auch die Kommentare, dass der Test als Selbsteinschätzungsinstrument bei den Studierenden sehr gut ankommt. Sie sind sehr zufrieden und schätzen das Angebot.

Beispielrückmeldung:

`Der Selbsteinschaetzungstest ist ein hervorragender Indikator des Standes der allgemeinen Mathematik zum Zeitpunkt des Starts.`

und melden eine Diversität zurück

`von zu einfach bis zu schwer, von zu viel bis zu wenig.`

Zitate aus Beispielrückmeldung:

`Die Fragen haben mir gezeigt, dass der Lernstoff meines Gymnasiums bei weitem nicht ausreicht!`

`Ich war etwas geschockt wie schlecht ich war, da ich eigendlich` [sic!] `eine 6 im Maturazeugnis habe, ich hier jedoch nur 7 Punkte erreichte XD`

2.6. Aftermath

Das arithmetische Mittel der erreichten Punkte liegt bei einem Maximum von 28 (ab Herbst 2013 von 29) jedes Jahr in einem ähnlichen Bereich. Die Tabelle zeigt das Mittel der Jahre 2010 bis 2015 für sieben Vorlesungen der Natur- und Ingenieurwissenschaften.

	1	2	3	4	5	6	7
	LifeSc	Umwelt	Chemie	BauIng	MaschIng	ElektIng	Informatik
2010	15	14	17	16	18	19	19
2011	15	15	18	18	20	20	19
2012	14	13	18	16	19	19	20
2013	12	12	15	15	16	18	17
2014	13	13	16	16	17	19	18
2015	12	12	13	15	18	18	18

3. Duales System: Vorlesungen und Übungen

Neben der qualitativen Entwicklung stellt auch das Wachstum der Studierendenzahl der letzen Jahre eine Herausforderung an die Mathematikausbildung dar. Mit vertretbaren Ressourcen ist ein ausgewogenes Betreuungverhältnis für eine intensive Begleitung der Studierenden im Basisjahr kaum möglich.

Nach F. Weinert muss Lernen sowohl sachsystematisch (Vorlesung) als auch situiert (Übung) erfolgen.

> *"Systematisch erworbenes Wissen – so die These – ist anders strukturiert, anders organisiert und anders abrufbar als es die meisten praktischen Anwendungssituationen erfordern. Prinzipiell verfügbares Wissen bleibt deshalb* **träge** *und* **ungenutzt***, obwohl man es eigentlich zur Lösung bestimmter Probleme braucht. Die Diskrepanz zwischen Lern- und Anwendungsbedingungen ist in der Regel sehr groß. Inzwischen lässt sich die wissenschaftlich fundierte Schlussfolgerung ziehen, dass Lernen sowohl sachsytematisch als auch situiert erfolgen muss.*
> *Mit anderen Worten: Neben einem wohl organisierten disziplinären Wissenserwerb bedarf es von Anfang an einer Nutzung des erworbenen Wissens in ... problemorientierten Kontexten. Die Förderung sowohl des situierten als auch des systematischen Lernens ist eine wesentliche Bedingung für den Erwerb von intelligentem, flexibel nutzbarem Wissen."* *(vgl. Weinert (1998), S. 24)*

Ein Beispiel für träges Wissen

In Vorlesung und Übung wurde erarbeitet und ausführlich erörtert:

Das lineare Gleichungssystem $Ax = 0$ hat genau dann nichttriviale Lösungen, wenn $\det(A) = 0$.

In der summative Prüfungsaufgabe:

> Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.
>
> a) Bestimmen Sie einen Vektor $x \neq 0$ sodass $Ax = 0$.
> b) Bestimmen Sie die Determinante von A.

folgte das ernüchterndes Resultat: Fast alle Kandidat/-innen berechnen $\det(A)$!

Zusammenfassung

Übungen ermöglichen, die situative Auseinandersetzungen mit dem Stoff und:

- gegenseitige Rückmeldungen.
- das neu gelernte Wissen zu sichern, Lernprozesse anzustossen und zu unterstützen.
- passives in aktives Wissen zu transformieren und die Studierenden mathematisch handlungsfähig zu machen.
- "kognitive Konflikte" zu erkennen und zu beseitigen und zu einem *besseren*, i.e. robusteren Verständnis zu führen.

Aber die inhärente Schwierigkeit bleibt, dass nicht alle Studierenden die gleichen Bedürfnisse haben.

3.1. Formatives und Summatives Assessment

In den Lehrveranstaltungen unterscheidet die Didaktik zwischen formativem und summativem Assessment (vgl. Dahinden (2014), Kapitel 3.2.2).

Formatives Assessment

Formatives Assessment ist im Rahmen einer formativen Leistungskontrolle (LK) Bestandteil des Lernprozesses und dabei meist unbenotet. Das Ergebnis einer formativen LK dient den Studierenden zur individuellen Leistungsbestimmung. Zentrales Element von formativen LK ist die Rückmeldung (Feedback). Die Qualität des Feedbacks beeinflusst dessen Effekt auf den Lernprozess (ebda.). Häufige Formen von formativen Leistungskontrollen sind Selbsttests, Übungsaufgaben und Besprechungen mit Lehr- und Assistenzpersonen.

Summatives Assessment

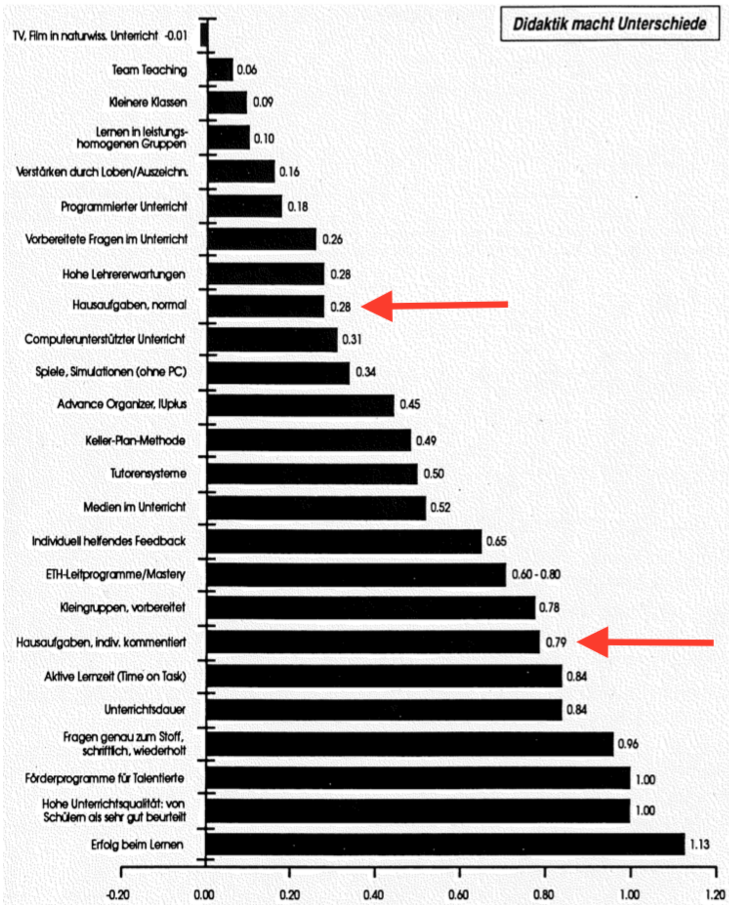
Die summative LK bildet den Abschluss einer Unterrichtssequenz und misst abschliessend die kognitive Leistung der Studierenden. Die Literatur bezeichnet dieses abschliessende Wissen als Learning Outcome (ebda.). Das Resultat wird in Form einer Note ausgedrückt. Rückmeldungen sind optional.

3.2. Mehrwert Rückmeldungen

Ein zentraler Mehrwert einer Lehrveranstaltung ist die individuelle Rückmeldung im Sinne eines formativen Assessment (vgl. Hattie (2013), S. 206ff.). Rückmeldungen beeinflussen die Leistung, die Motivation und steigern das Bewusstsein über die Selbstwirksamkeit des Lernens (vgl. Dahinden (2014), Kapitel 3.4). Einige Studien zeigen gar, dass Feedbacks zu den effektivsten didaktischen Mitteln gehören, um die Leistung von Studierenden zu steigern (ebda.). Auch schwächere Studierende mit weniger Vorwissen können vom Feedback profitieren (ebda.).
Dies hilft uns bei einer Homogenisierung und Unterstützung der Studierenden im Basisjahr.

Aus der empirischen Unterrichtsforschung

Ein Vergleich unterschiedlicher Methoden und Massnahmen auf einer $[-0.2, 1.2]$-Effektstärkenskala zeigt die Wirkung auf den Schul- und Studienerfolg (vgl. Frey, Frey-Eiling (2004), Kapitel 1). Danach führt die **individuelle** Kommentierung von Hausübungen zu einer Verbesserung von $+0.5$ im Gegensatz zu einfachen Hausübungen.



4. Das MC-Format

Jede Woche betreut ein/e Lehrassistierende/r in Präsenzübungen maximal 25 Studierende mit Diskussion und Detailkorrekturen der Übungsaufgaben und weiteren Fragen. Die gewachsene Studierendenzahl erlaubt bei umsichtigen Ressourcen weniger bis kaum noch die wichtige individuelle Rückmeldung.

Die virtuellen Zusatzlernangebote ermöglichen Ergänzungen und

- bieten den Studierenden während des eigenständigen Lernens, adaptive sowie interaktive Instrumente zur Selbsteinschätzung und zum wiederholten Üben und Vertiefen ihrer Kompetenzen,
- erlauben ihnen, ihren Wissensstand und das Verständnis des Lerninhalts nach Bedarf zu überprüfen, zu vertiefen und Defizite zu beheben.

Bereits das elementare Eingabeformat Multiple-Choice generiert für Studierende sowie Dozierende einen Mehrwert.

4.1. Mehrwert für Studierende

Die MC-Aufgaben regen die Studierenden an, mathematische Probleme eigenständig zu lösen (vgl. Caspar, Miller (2012), Kapitel 1.1.1) und bei Bedarf kreative Lösungswege zu generieren (ebda.). Die MC-Antwortvorgaben verlangen aber von den Studierenden auch eine präzisere Bearbeitung.

Als zentraler Mehrwert gilt, dass die Online MC-Aufgaben den Studierenden ein unmittelbares und eindeutiges Feedback auf ihre Lernleistungen generieren. Sie können ihren Leistungsstand einschätzen und das weitere Lernen planen (ebda.).

Die unmittelbaren Rückmeldungen unterstützen den Lernprozess durch eine partielle Individualisierung und sind so differenziert, dass die Studierenden erkennen können, wo Vertiefungs- und Nachholbedarf bestehen (ebda.). Diese formativen Prüfungen zeigen, wie die Leistungen im Vergleich zu den Mitstudierenden stehen, was in Hinblick auf die summativen Prüfungen bedeutsam ist.

4.2. Mehrwert für Dozierende

Die Dozierenden erhalten detaillierte statistische Auswertungen über die Kompetenzen der Studierenden. Sie ersehen, bei welchen Aufgaben typische und/oder systematische Probleme auftreten und können anhand dieser Kenntnisse die Lehrveranstaltung adaptiv planen, modifizieren und bei Bedarf verbessern. Unklarheiten, fehlende Denkschritte usw. können aufgegriffen und geklärt werden (vgl. Caspar, Miller (2012), Kapitel 1.1.2).

Mit MC als Online-Angebot erreichen Dozierende eine höhere Effizienz: Der umsichtig geplante und konzipierte Einsatz entlastet sie von Routinearbeiten (ebda.). Es entsteht freie Zeit für anspruchsvollere Lehrtätigkeiten und die Betreuung der Studierenden. Dies erlaubt die Umsetzung neuer Lehrformen und Bearbeitung kognitiv anspruchsvollerer und komplexerer Themen.

4.3. Implizites Wissen

Das geschlossene Aufgabenformat einer MC-Frage überprüft zunächst nur das Ergebnis, nicht aber die Leistung der/s Studierenden, welche sie/er für die Lösung aufwänden musste. Wir können bei einer einzelnen Fragen aufgrund der Antwort nicht auf Schwierigkeiten und den Lösungsweg schliessen und die Frage nicht einordnen. Die Kontrolle über das implizite Wissen der Studierenden fällt weg. Es ist kein oder nur aufwändig ein iterativer Lehr/Lernprozess nach einem Don't say, ask! - Prinzip von P. Halmos möglich:

> *"The hardest part of teaching by challenging is to keep your mouth shut, to hold back.*
> *Don't replace the wrong A by the right B, but ask, "where did A come from?"*
> *Keep asking "Is that right? Are you sure?" Don't say 'no'; ask 'why?'."*
> *(Halmos (1985), S. 272)*

Folgen von Fragen und deren Antwortschemata könnten unter Umständen helfen, das implizite Wissen der Studierenden aufzudecken.

4.4. Erstellen einer Multiple-Choice-Aufgabe

Formal-technisch unterliegen MC-Fragen einem einfachen Format. Herausfordernd und anspruchsvoll sind das Finden einer Aufgabe auf dem angestrebten kognitiven Niveau mit geeigneten Distraktoren, ausführlichen Rückmeldungen und der Differenzierung für einen formativen oder summativen Einsatz.

Die Aufgaben müssen so konzipiert, dass sie ein kontinuierliches Repetieren und Überprüfen der Lerninhalte mit zunehmendem Schwierigkeitsgrad erlauben. Die Untersuchung der praxiserprobten Aufgaben kann uns bei der Erstellung helfen.

4.5. Einsatzszenario Einschätzung

Wir unterscheiden die Einsatzszenarien von MC-Fragen (vgl. Caspar, Miller (2012), Kapitel 2.2):

$$\underbrace{\text{Prüfung}}_{\text{Summativ}} - \underbrace{\boxed{\text{Einschätzung}} - \text{Quiz}}_{\text{Formativ}}$$

Prüfung Einmaliges Abschicken **aller** Aufgaben bis zum Einsendeschluss. Die Studierenden erhalten Lösung, Rückmeldung und Auswertung nach dem Einsendeschluss.

Einschätzung Einmaliges oder wiederholtes Abschicken **aller** Aufgaben bis zum Einsendeschluss. Die Studierenden erhalten Lösung, Rückmeldung und Auswertung nach dem Absenden.

Quiz Hier werden die **einzelnen** Aufgaben mit einer sofortigen Rückmeldung wiederholt abgeschickt.

4.6. Raten der Lösung

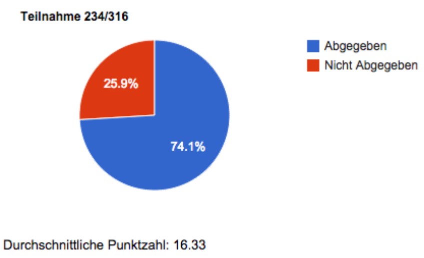
Der geschlossene MC-Aufgabentyp erlaubt zunächst nur das Ergebnis zu überprüfen, nicht aber die Leistung und das implizite Wissen der/s Studierenden für die Lösungsfindung. Hinzu kommen Raten oder zufällige Auswählen einer MC-Antwort. Diese bereiten weitere Schwierigkeiten, auf die Lösungsstrategien und allfällige Probleme der Studierenden zu schliessen.
Das Szenario Einschätzung bietet bei jeder MC-Frage die Option “Weiss ich nicht.” und die Empfehlung, diese auch tatsächlich zu wählen. Damit wird das Raten eingeschränkt, und es besteht für die Studierenden keine nachvollziehbare Motivation, zu raten oder die Fragen willkürlich zufällig zu beantworten. Damit erhalten wir zuverlässigere Daten.
Mit der Einführung von “Weiss ich nicht.” steht uns auch die Messung der Selbsteinschätzung der Studierenden zur Verfügung. Wer wählt die Option bei welchen Aufgaben? Welche Studierenden schätzen sich angemessen ein? Bestimmt dies das Vorwissen oder andere Parameter? Gibt es einen Gender-Gap? Eine Untersuchung dieser Fragen wird vorbereitet.

4.7. Datenerhebung

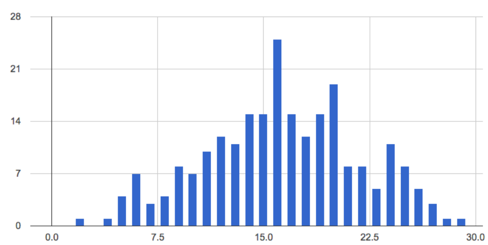
Jede Semesterwoche bearbeiten die Studierenden eine Serie mit schriftlichen und Online-MC-Aufgaben.

Die Auswertung nach Einsendeschluss einer Serie umfasst eine detaillierte Statistik der Abgaben mit

- einer Verteilung der Gesamtpunktzahl, auch pro Übungsgruppe.
- einer prozentuale Verteilung auf die Antworten einzelner Aufgaben, auch für die eigene Gruppe.
- einer Einordnung der eigenen Leistung insgesamt und relativ zu meiner eigenen Gruppe.



Beteiligung	$\frac{234}{316}$	74%
Erreichbare Punktzahl	29	
Maximal erreichte Punktzahl	29	
Minimal erreichte Punktzahl	2	
Arithmetisches Mittel	16.33	



Genau die korrekten Antworten: ca. 56% / 58% --

Die Matrix $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ hat die Eigenwerte .

	Ca. 13%/14%	◯	-1
√	Ca. 75%/85%	◯	1
√	Ca. 81%/76%	◯	i
√	Ca. 78%/66%	◯	$-i$
	Ca. 0%/2%	◯	$1-i$

5. Item-Analyse

Für einen weitergehenden Beitrag zur Qualitätssicherung und -entwicklung in der mathematischen Grundausbildung suchen wir eine Verbesserung der Aufgaben und Einsatz von adaptiven Folgen und Empfehlungen für weiteres Lernen.

Wir möchten durch eine quantitative Datenanalyse zum Beispiel eine Triage-Messung erhalten für leichte – schwierige Aufgabe zur Unterstützung schwacher – starker Studierender.

Zur Systematisierung dieser Messung sortieren wir die Aufgaben in Taxonomien als Basis für eine zuverlässige Messung der kognitiven Leistungen.

5.1. Taxonomie Mathematischer (MC-)Aufgaben

Beim Stellen von Übungen fragen wir uns zum Beispiel:

- Welche Leistung braucht es beim Lösen einer Aufgabe?
- Was ist ein geeignetes Verhältnis von Aufgaben in einer Serie?

Mit Hilfe einer Sortierung der Aufgaben unterscheiden wir diese Leistung. Eine inhaltsorientierte Klassifizierung deuten wir in (Caspar, Miller (2012), Kapitel 2.1) an

Typ	**Beispiel**
Repetition	”Unter welchen Voraussetzungen gilt die Aussage?”
Rechnung	”Bestimmen Sie das korrekte Resultat!”
Verständnis	”Welche Eigenschaften kann ein Objekt haben?”
Transfer	”Welche Aussagen sind nicht äquivalent?”

Dies dient zum Einen einer vereinfachten Konstruktion und Standardisierung der MC-Aufgaben, kommt aber auch einem Bedürfnis der Studierenden nach, die unterschiedlichen Lernziele und Eigenschaften einer Aufgabe zu erkennen. Dies hatten sie in Evaluationen zurückgemeldet.

Diese Typisierung nach Inhalten entwickeln wir nun zu einer hierarchischen Taxonomie weiter, welche sich an der Literatur orientiert (vgl. Dahinden (2014), Kapitel 2.3).

Als Basis für eine zuverlässige Messung der kognitiven Leistungen ermöglichen Taxonomien Fragen zum Prüfungserfolg (Korrelieren Stufen mit Prüfungserfolg?), zur Schwierigkeit (Spiegelt die Erfolgsquote die Kognitionsstufe wieder?) oder zu den Studiengängen/Lernzielen (Zeigen die gleichen Fragen andere Auswertungen?)

Die bekannten Taxonomien nehmen an, dass jedes Lehrgebiet, “from art history to zoology” (vgl. Fuller et al. (2007), Kapitel 1.1) die jeweilige Hierarchie respektiert. Oft sind diese mühsam anzuwenden. Eine Einordnung in Stufen ist schwierig oder zumindest uneindeutig. Auch kann die Hierarchie nicht konsequent umgesetzt werden.

Wir skizzieren hier vier Taxonomien.

ETHZ-S21t

Um die Fragen des S21t nicht allein inhaltsorientiert zu entwickeln, hatten wir etwas handgestrickt die Taxonomie KOG1 – KOG2 – KOG3 in 2.3 vorgeschlagen.

In unseren Auswertungen ist der Anteil richtiger Antworten in KOG1 und KOG2 jeweils etwa vergleichbar, KOG3 wird deutlich schlechter beantwortet. Es gibt nur eine relativ geringe Korrelation mit den Notenschnitten in der Prüfung. Diese Korrelation ist in den drei Bereichen ähnlich.

CH-Kanon

Eine gesamtschweizerische Arbeitsgruppe hat den “Katalog Grundkenntnisse der Mathematik” von 1997 überarbeitet, Link zur Webseite http://www.math.ch/kanon/ .

Er bezieht sich aufs Grundlagenfach Mathematik an den Gymnasien und befasst sich inhaltlich mit der Situation an der Schnittstelle Gymnasium / Universitäre Hochschulen. Der Katalog umreisst mathematische Kenntnisse und Fähigkeiten der Maturanden/innen und damit das mathematische Wissen und Können der Studienanfängerinnen und -anfänger auf der tertiären Stufe.

Der Kanon unterteilt ein Gebiet in Semantik – Syntax – Exploration.

Dabei ist Semantik gleich verstehensorientiertes inhaltliches Wissen, Syntax gleich verfahrensorientierte, algorithmische Fertigkeiten und Exploration gleich verstehensorientierte Erkundung/Vertiefung.

Zum Beispiel für die Grundlagen der Differentialrechnung

Semantik	Syntax	Exploration
• Mass für die Veränderung: Mittlere und momentane zeitliche Änderungsrate mittels Anwendungen • Linearisierung, Tangente, Steigung (z.B. Monotonie) • Extremal- und Wendepunkte • Optimierungsanwendungen	• Differenzenquotient • Differentialquotient / Ableitung • Ableitungsfunktion • Grundeigenschaften der Ableitung (Linearität) • Ableitung der Grundfunktionen • Ableitung von $f(ax+b)$ • Extremalprobleme	• Lineare Approximation, Newtonverfahren • Anwendungen in den Natur- und Sozialwissenschaften sowie der Ökonomie (Kinematik, Wachstum und Zerfall) • Ableitungskalkül (Produkt-, Quotienten-, Kettenregel)

Klassisch 1-Dimensional

Die populärste Taxonomie geht auf Bloom (vgl. Bloom (1956)) zurück. Er entwickelt eine eindimensionale hierarchische Struktur mit Kognitionsstufen K1 bis K6 in einer Kognitionsprozessdimension.

$$\underbrace{\underbrace{\text{Erinnern}}_{\text{K1}} – \text{Verstehen} – \underbrace{\text{Anwenden – Analysieren – Bewerten – Erzeugen}}_{\text{K3+}}}_{\text{Was muss ich verarbeiten? Kognition}}$$

Die Hochschulebene strebt vor allem die Stufen ab K3 an, welche mit K3+ zusammengefasst werden. Der Übergang von der Inhalts- zur Kompetenzvermittlung erfolgt von in der Regel von K2 zu K3 (vgl. Dahinden (2014), Kapitel 2.5).

Wissens- oder Handlungsdimension

Anderson et al. schlagen in Anderson et al. (2001) eine zweidimensionale Lösung mit einer zusätzlichen dreistufigen **Wissensdimension** vor. Dabei beschreibt Faktenwissen nicht unbedingt eine geringere Leistung als Konzeptwissen und dies nicht unbedingt als Vorgehenswissen. In Fuller et al. (2007) findet sich der Vorschlag, die Kognitionsstufen nach Bloom zu unterteilen und eine Handlungsdimension einzuführen.

Kognition (Bloom) Was muss ich verarbeiten?	**Neue 2. Dimension** Was muss ich **wissen** / tun?
Bewerten Analysieren Verstehen Erinnern	**Fakten Konzept Vorgehen** Nichts Anwenden Erzeugen Oder doch anders?

5.2. Fehlermatrix/-verteilung

Eine Fehlermatrix visualisiert[5] die Fehlerverteilung (vgl. Dahinden (2014), Anhang A.5):

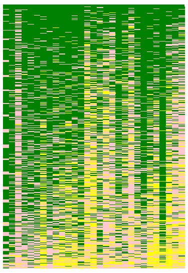
Binäre Verteilung

Eine Zeile zeigt die Resultate eines/r Studierenden, nach Erfolgsquoten sortiert: von den Erfolgreichsten abwärts.

Eine Spalte zeigt die Fehlerverteilung einer Aufgabe. Grüne Bereiche entsprechen korrekten Antworten, pinkfarbende falschen oder fehlenden Antworten.

Die Matrix gibt einen Überblick: Pinklastige Spalten deuten auf schwierige Fragen hin, grünlastige auf leichte. Zudem sollten sich die pink Felder nach unten verdichten.

Ternäre Verteilung



Seit 2013 bietet jede MC-Frage die Option "Weiss ich nicht." und die Empfehlung, diese auch tatsächlich zu wählen. Damit wird das Raten eingeschränkt, und es besteht für die Studierenden keine nachvollziehbare Motivation zu raten, oder die Fragen willkürlich zufällig zu beantworten.

Hier entsprechen grüne Bereiche wieder korrekten Antworten, pinkfarbende falschen/fehlenden Antworten und gelbe **"Weiss nicht"-Antworten**.

Die Abbildungen links zeigen die bi/ternäre S21t-Verteilungen im Herbst 2013.

5.3. Erfolgsquote

Wir nehmen zunächst an, dass eine Frage nicht durch raten oder zufälliges auswählen gelöst wird. Dann ist die Erfolgsquote oder Schwierigkeitsindex der i-ten Frage:

$$P_i = \frac{\text{Zahl der korrekten Antworten}}{\text{Zahl aller \textbf{(abgegebenen)} Antworten}}$$

mit den Richtwerten (vgl. Dahinden (2014), Tabelle 3.6).

$0.2 < P_i < 0.8$ angestrebte Minimalstreuung je nach Differenzierungsziel,

$P_i < 0.2$ bei einer eher schwierigen Frage,

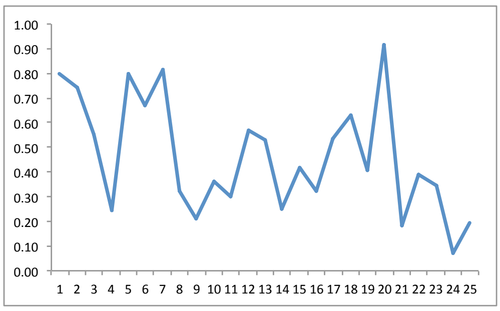
$P_i \geq 0.95$ bei einer einfachen (Eisbrecher-)Frage.

Optimierung Erfolgskurve

In der Regel erfolgt der Einstieg mit einfachen (Eisbrecher-) Fragen.

Wie geht es dann weiter? Gibt es beim Verlauf der Erfolgsquoten $i \mapsto P_i$ eine ideale Erfolgskurve zur optimalen Unterstützung von Lernerfolg/-motivation? Wann ist $P_{i+1} \leq P_i$, wann ein Wechsel $P_{i+1} \geq P_i$ angezeigt, welche Funktion Φ mit $P_{i+1} = \Phi(P_i)$ unterstützt optimal das Lernen, die Motivation oder das Verstehen?

Erfolgskurve der S21t in 2009



[5] Für die farbliche Unterscheidung bitte die online pdf-Version des Artikels beachten.

5.4. Trennschärfe

Die Trennschärfe[6] einer Frage gibt ein Mass, wie gut diese Frage die Rangliste des Gesamtergebnisses widerspiegelt. Sie ist eine Zahl r mit $-1 < r < 1$ und stellt die Verteilung der korrekten Antworten der Studierenden bei der jeweiligen Frage in Bezug zu der Punktverteilung des Schlussergebnisses.

Wir schreiben wieder das Testergebnis in eine Matrix. In jeder Zeile stehen die Punkte eines Studenten.

$\downarrow$ i-te Frage

Jede Spalte (= Frage) liefert die Verteilung der richtigen (grün) / falschen (pink) Antworten der Studierenden. Diese steht in Beziehung zur Spalte mit der Summe an richtigen Antworten einer/s Studierenden.

Die Trennschärfe r_i der i-ten Frage wird unterschiedlich definiert. Die populärste Variante ist der Korrelationskoeffizienten $\text{cor}(X,Y)$ mit $X =$ der Punktzahl der i-ten Aufgabe und $Y =$ der erreichten Punkte eines Studenten.

- Eine hohe Trennschärfe besagt, dass insgesamt gute Studierende diese Frage richtig und schwächere falsch oder nicht beantwortet haben.

  Bei $0.3 < r_i$ haben wir eine gute Frage, eine Frage mit $0.2 < r_i \leq 0.3$ ist noch akzeptabel.

  Eine Trennschärfe unter der empfohlenen Grenze von 0.3 impliziert nicht automatisch ein Ersetzen des Items – vielmehr erhalten wir Informationen, warum diese Frage nicht gut mit dem Gesamtergebnis korreliert: Gibt es ein breites Missverständnis oder lokal in einer Gruppe von Studierenden?

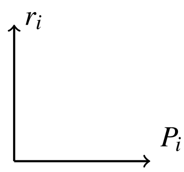
- Bei einer Trennschärfe um 0 haben ungefähr gleich viele gute wie schwächere Studierende korrekt geantwortet.

- Bei einer negativen Trennschärfe handelt es sich um eine "Fangfrage", welche mehrheitlich von schwachen Studierenden korrekt und von stärkeren Studierenden falsch beantwortet wurde. Sie verfälscht das Ergebnis. In den Szenarien Prüfung und Einschätzung ist sie eher ungeeignet.

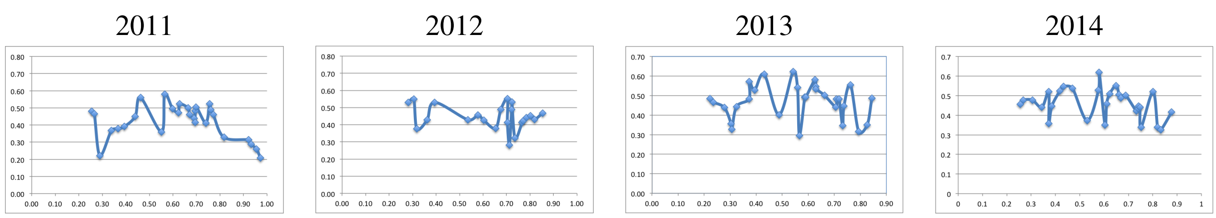
  Bewusst eingesetzt, kann sie didaktisch sinnvoll sein. Die Lernenden müssen dazu eine ausführliche Rückmeldung erhalten.

5.5. Trennschärfe vs. Erfolg

Zwischen Erfolgsquote und Trennschärfe besteht die Abhängigkeit, bei mittlerer Erfolgsquote höhere Trennschärfen zu erwarten als bei schwereren oder leichteren Aufgaben. Idealerweise ähnelt die Kurve im (P,r)-Koordinatensystem einer nach unten geöffneten gestreckten Parabel.



$r(P)$-Kurven für den S21t von 2011 bis 2014

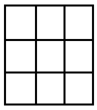


6 Wir folgen wieder im Wesentlichen Dahinden (2014), Kapitel 3.3.4.

Ausreisseraufgaben

Der parabolische Verlauf lässt sich jeweils erahnen. Interessant ist, dass in 2013 und 2014 jeweils die gleichen Aufgaben bei einer mittleren Erfolgsquote $P \approx 0.6$ nach unten ausreissen.
Aus der Kombinatorik

Sara malt die 9 Felder einer Zeichnung mit Farbstiften an. Sie besitzt Stifte in 12 unterschiedlichen Farben. Sara malt genau 3 Felder gelb an und die weiteren Felder jeweils mit einer beliebigen der anderen Farben.



Wie viele Möglichkeiten hat sie dafür? a) $\binom{9}{3}$ b) 11^6 c) $\binom{9}{3} \cdot 11^6$ d) $9!$ e) $12^9 - 3$

Zur Geometrischen Reihe

$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \ldots$ ist gleich ... a) $\frac{5}{8}$. b) $\frac{2}{3}$. c) $\frac{11}{16}$. d) $\frac{3}{2}$. e) ∞.

Hier könnte die tiefe Trennschärfe ein Hinweis sein, dass auch schwächere Studierende diese Aufgaben richtig beantworten können.

Binär versus Ternär

Erfolgsquote und Trennschärfe in der binären und der ternären Verteilung unterscheiden sich. Tendenziell wird $P_{\text{binär}} \leq P_{\text{ternär}}$ gelten, denn für die Berechnung von $P_{\text{ternär}}$ ist die Zahl der Abgaben kleiner.
Für die Trennschärfe ist $r_{\text{binär}} > r_{\text{ternär}}$, wenn sich die gelben Felder in der Matrix in 5.2 nicht homogen verteilen und sich im unteren Bereich verdichten. Nehmen wir diese Ergebnisse aus der Berechnung, erhalten wir ein schwächere Korrelation.

6. Beispielaufgaben aus S21t und deren Streuung

In Akveld, Caspar (2014) diskutieren wir ausgewählte Aufgaben des S21t. Dies führen wir hier fort und betten die Beobachtungen in die Item-Analyse ein. Das Ziel ist eine Klassifizierung die Aufgaben in eine Triage mit einem oberen / unteren Drittel und einer unentschiedenen Mitte.

Wir geben jeweils an: $P = P_{\text{ternär}} =$ Erfolgsquote, $NA =$ Zahl "Weiss ich nicht", $r = r_{\text{binär}} =$ Trennschärfe und an einigen Stellen auch auffällige (Auswahlquoten). Angegeben sind die Daten aus 2013, im Folgejahr lagen sie im ähnlichen Bereich.

6.1. Oberes Drittel

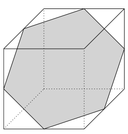
Im oberen Drittel vermuten wir die eher stärken Studierenden oder Studierende, welche eine solide Vorbildung mitbringen. Sie lösen Aufgaben mit einer Erfolgsquote P von ca. $P \leq 0.33$.

Die beiden folgenden Aufgaben zeigen ein tiefes P mit gutem r und trennen daher zuverlässig die sehr guten Studierenden von den guten.

Raumvorstellung

Die Schnittmenge eines Würfels mit einer Ebene sei ein Vieleck. Bestimmen Sie die maximale Anzahl von Ecken dieses Vielecks.

a) 3 b) 4 (0.56) c) 6 d) 8



$P \approx 0.32$, $NA \approx 0.05$, $r \approx 0.32$

Grenzwerte

Der Grenzwert $\lim_{h\to 0} \frac{\sqrt{2+h}-\sqrt{2}}{h}$ ist gleich ... a) 0. (0.45) b) $\frac{1}{2\sqrt{2}}$. c) $\frac{1}{2}$. d) $\frac{1}{\sqrt{2}}$. e) ∞(0.2).
$P \approx 0.27$, $NA \approx 0.20$, $r \approx 0.51$

In zwei weiteren Aufgaben über Grenzwerte $\lim_{n\to\infty} \frac{2n^3-1}{10n^3+n+21} = \ldots$ und $1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\frac{1}{16}-\ldots = \ldots$ sind die Studierenden wesentlich erfolgreicher, jeweils liegt jeweils P im oberen Bereich.

Wurzel 36

Die Wurzel aus 36 ... a) Gibt es nicht. b) Ist gleich ± 6. (0.65) c) Ist gleich 6. d) Ist gleich -6. e) Keine der obigen Antworten ist richtig.
$P \approx 0.34$, $NA \approx 0.0$, $r \approx 0.17$

Diese Frage hatten wir nur im Herbst 2009 gestellt. Sie scheint einfach, aber nur 34% antworteten richtig. Gleichzeitig ist $r \approx 0.17 < 0.3$ eine weniger gute Trennschärfe. Die Erfolgsquote bei dieser Frage korreliert nicht mit den Gesamtergebnis: Auch insgesamt erfolgreiche Teilnehmende wussten die Antwort nicht, und vice versa. Aufgrund des erratischen Antwortverhaltens der Studierenden haben wir diese Aufgabe in den folgenden Jahren nicht mehr gestellt.

Es wählten 65% die zweite Variante, und glaubten offenbar, dass die (Quadrat-)Wurzel nicht eindeutig definiert ist. Unsere Vermutung ist, dass die Schüler/innen das Lösen der Gleichung $x^2 = 36$ mit dem Wurzelziehen durcheinanderbringen. Die Lösungen der Gleichung sind $x_{1,2} = \pm\sqrt{36}$, aber das $\pm$ - Zeichen ist nicht in der Wurzel inbegriffen.

6.2. Unteres Drittel

Im unteren Drittel vermuten wir die eher schwächeren Studierenden oder Studierende mit weniger Vorbildung. Sie können Aufgaben mit ca. $P \geq 0.66$ nicht lösen.

Differentialrechnung

Sei f die Funktion mit $f(x) = e^{2x}$. Wie lautet die Gleichung der Ableitung?
a) $f'(x) = 2xe^{2x-1}$ b) $f'(x) = \frac{1}{2}e^{2x}$ d) $f'(x) = 2e^{2x}$ e) $f'(x) = e^{2x}$
$P \approx 0.8$, $NA \approx 0.2$, $r \approx 0.55$

Diese Aufgabe erkennt zuverlässig schwächere Studierende.

Ohne Kommentar dokumentieren wir noch weitere Aufgaben.

Algebra

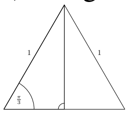
Welche der folgenden Rechenregeln stimmt für alle positiven reellen Zahlen a und b?
a) $\frac{1}{a+b} = \frac{1}{a}+\frac{1}{b}$ b) $\sqrt{a+b} = \sqrt{a}+\sqrt{b}$ c) $(a+b)(c+d) = ac+bd$ d) $\ln(a+b) = \ln(a)+\ln(b)$, (0.2) e) Keine. $P \approx 0.72$, $NA \approx 0.02$, $r \approx 0.46$

Welcher der folgenden Ausdrücke ist für $a,b > 0$ gleich $\ln(a^4b^2) - \ln(a^2b^{-2})$
a) $6\ln(a)$ b) $2\ln(a) - 4\ln(b)$ c) $\frac{\ln(a^2b)}{\ln(ab^{-1})}$ d) $\ln(a^2b^4)$
$P \approx 0.73$, $NA \approx 0.23$, $r \approx 0.39$

Trigonometrie

Bestimmen Sie $\sin\left(\frac{\pi}{3}\right)$.

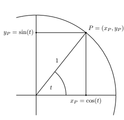
a) 0 b) $\frac{1}{2}$ c) $\frac{\sqrt{2}}{2}$ d) $\frac{\sqrt{3}}{2}$ e) 1
f) Das geht nur mit einem Taschenrechner. (0.18)

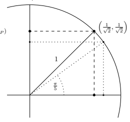


$P \approx 0.67, NA \approx 0.12, r \approx 0.45$

Für welches n ist $\cos\left(\frac{\pi}{n}\right) > \sin\left(\frac{\pi}{n}\right)$?

a) $n = 2$ b) $n = 3$ c) $n = 4$ d) $n = 5$ e) Das geht nur mit einem Taschenrechner. (0.07)





$P \approx 0.74, NA \approx 0.15, r \approx 0.43$

6.3. Unentschiedene Mitte

1. Das Integral $\int_0^2 3x^2\,dx$ ist gleich ...

   a) $\frac{4}{3}$ b) 2 c) $\frac{8}{3}$ d) 4 e) 8 $P \approx 0.94, NA \approx 0.1, r \approx 0.28$

2. Das Integral $\int_0^1 e^{-2t}\,dt$ ist gleich ...

   a) $1 - \frac{1}{e^2}$ b) $\frac{1}{2e^2}$ c) $\frac{1}{2} - \frac{1}{e^2}$ d) $1 - \frac{1}{2e^2}$ e) $\frac{1}{2} - \frac{1}{2e^2}$ $P \approx 0.55, NA \approx 0.32, r \approx 0.45$

3. Das Integral $\int_{-1}^1 |t|\,dt$ ist gleich ...

   a) 0 (0.43) b) 1 c) 2 d) 4 $P \approx 0.44, NA \approx 0.15, r \approx 0.44$

7. Zusammenfassung: Fazit – Ausblick

Webbasierte Systeme erlauben eine systematische Erfassung von Studierendenleistungen. Diese können mit einfachen Analysemethoden ausgewertet und beurteilt werden. Diese Messungen helfen uns, den Unterricht anzupassen und allenfalls zu verbessern.

Wir haben dies hier an den Aufgaben eines Selbsteinschätzungsstests über Schulmathematik an Beispielen demonstriert.

Der nächste Schritt konzentriert sich auf die Aufgaben aus dem ersten Studienjahr, welche in der Tendenz auf einer höheren Kognitionsstufen zu bearbeiten sind. Diese quantitative Item-Analyse könnte weitere Hinweise liefern für

- eine angemessene Taxonomie mathematischer Aufgaben (vgl. 5.1 oben),
- eine optimale Lern- oder Motivationskurve (vgl. 5.3 oben),
- ein Verständnis der Lernstrategien unterschiedlicher Gruppen (Gender) (vgl. 4.6 oben).

Literatur

Meike Akveld, Alexander Caspar. *Was ist die Wurzel aus* 36 *?* VSMP Bulletin 125 (2014).

Lorin Anderson et al. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* Longman (2001).

David Paul Ausubel. *Educational Psychology: A Cognitive View.* Holt, Rinehart and Winston. New York (1968).

Benjamin S. Bloom. *Taxonomy of educational objectives; the classification of educational goals (1st ed.).* Longmans, Green. New York (1956).

Alexander Caspar, Damian Miller. *MC-LaTeX-Weblikationen: Online-Multiple-Choice-Aufgaben in der mathematischen Grundausbildung der ETH Zürich.* Digitale Medien - Werkzeuge für exzellente Forschung und Lehre 393 – 400. Waxmann. Münster (2012).

Markus Dahinden. *Designprinzipien und Evaluation eines reliablen CBA-Systems zur Erhebung valider Leistungsdaten.*. DOI, Dissertation ETH Zürich (2014).

Karl Frey, Angela Frey-Eiling. *Allgemeine Didaktik: Arbeitsunterlagen zur Vorlesung.* ETH Zürich (2004).

Ursula Fuller et al. *Developing a Computer Science-specific Learning Taxonomy.* SIGCSE Bull. 39/4, 152 – 170, (2007).

Paul R. Halmos. *I want to be a Mathematican. An Automathography.* Springer. Berlin (1985).

John Hattie. *Lernen sichtbar machen*. Schneider-Verlag Hohengehren. Baltmannweiler (2013).

Michael Kerres. *Multimediale und telemediale Lernumgebungen. Konzeption und Entwicklung.* Oldenbourg Verlag. München, Wien (2001).

Michael Kerres, Claudia de Witt. Pragmatismus als theoretische Grundlage für die Konzeption von eLearning. In: Mayer, H. O. & Treichel, D. (Hrsg.) *Handlungsorientiertes Lernen und eLearning.* Oldenbourg Wissenschaftsverlag. München (2004).

Helmut M. Niegemann. *Kompendium E-Learning.* Springer. Berlin (2004).

Gabi Reinmann. Die vergessenen Weggefährten des Lernens: Emotionen beim eLearning. In: Mayer, H. O. & Treichel, D. (Hrsg.) *Handlungsorientiertes Lernen und eLearning.* Oldenbourg Wissenschaftsverlag. München (2004).

Franz E. Weinert. *Lehrerkompetenz als Schlüssel der inneren Schulreforn.* Schulreport 2/98. (1998).