

Prüfungen auswerten mit dem Niveaustufenmodell

Die Klausur ist geschrieben, die Präsentationen wurden gehalten, die Hausarbeiten sind abgegeben. Sie haben ein Semester lang mit den Studierenden gearbeitet und können nun an den Ergebnissen der Prüfung ablesen, wie erfolgreich die Lernprozesse waren. Was Sie dafür brauchen, ist ein verlässliches Instrument zur Bewertung der Prüfungsleistungen: Woran erkennen Sie eine Leistung der Notenstufe 1? Welche Leistungen haben Studierende erbracht, die gerade noch bestanden haben? Je konkreter Sie qualitativ beschreiben können, welche Kompetenzen Ihre Studierenden auf welchem Niveau erworben haben, desto transparenter, objektiver und fairer wird Ihre Bewertung sein.

Prüfungs-Auswertung für verschiedene taxonomische Stufen
Ein Beispiel
Werkzeuge identifizieren
Maximal-, Mindest- und Regelstandard festlegen
Literatur

Verfasser*innen:

↗[Michael Hagen, Birgit Szczyrba, Antonia Wunderlich \(2015\),](#)



Dieser Steckbrief ist lizenziert
unter einer ↗[Creative Commons](#)
[Namensnennung 4.0 International Lizenz.](#)

Die in diesem Steckbrief beschriebene
Methodik zur Formulierung von Learning
Outcomes wurde entwickelt von
Prof. Dr. Dr. Oliver Reis
(↗oliver.reis@uni-paderborn.de).

*Wir freuen uns über Fragen oder
Anregungen zu unseren Steckbriefen!*

Ansprechpartner*in:

Team Hochschuldidaktik

Zentrum für Lehrentwicklung

T: 0221-8275 3820

E: hochschuldidaktik@th-koeln.de

ZLE

Zentrum für

Lehrentwicklung

Technology

Arts Sciences

TH Köln

Prüfungs-Auswertung für verschiedene taxonomische Stufen

Für [Prüfungen auf den unteren Taxonomiestufen](#) fällt die Auswertung meist leicht: So ist z.B. bei der Definition eines Fachbegriffes, einer Handlung auf Taxonomiestufe 2 „Verstehen“, lediglich relevant, ob sie a) korrekt und b) vollständig wiedergegeben wurde. Dies lässt sich über die Identifikation von relevanten Aspekten („Welche Elemente der Definition wurden genannt?“) und ein einfaches Richtig-Falsch-Schema („Wurden diese Elemente korrekt verwendet?“) bewerten. Oder: beim Anwenden einer Formel lässt sich recht einfach erkennen, ob sie korrekt angewendet oder ob Rechenschritte ausgelassen wurden. Solchen Leistungen werden Sie gerecht, wenn Sie etwa für jeden relevanten und richtigen Aspekt in einer Definition oder jedes korrekt errechnete Ergebnis einen Punkt vergeben.

Bewertungssysteme sollten aber auch in der Lage sein, komplexere Leistungen objektiv zu benoten. Diese taxonomisch anspruchsvolleren Prüfungen können nicht genauso bewertet werden wie einfache Aufgaben auf niedrigeren Taxonomiestufen, weil die Maßstäbe „richtig/falsch“ oder „vollständig“ zwar bei ihrer Bearbeitung eine Rolle spielen, aber nicht genügen. Für die Analyse einer Unternehmensstrategie im Bereich Marketing z.B. brauchen Sie zusätzlich Kriterien wie „Angemessenheit“, „Kreativität“, „Begründetheit“ oder „Datenbasiertheit“. Um komplexe Leistungen zu bewerten, können also nicht einfach Punkte vergeben werden, sonst reduziert Ihre Notengebung den Anspruch, den Sie im [Learning Outcome](#) formuliert haben, steuert das Lernen in eine falsche Richtung (vgl. [Steckbrief Constructive Alignment](#)) und wird für die Studierenden intransparent. Der Ansturm auf Ihre Sprechstunde („Warum hab' ich nur eine Drei?“) ist dann vorprogrammiert.

Wie finden Sie eine Antwort darauf, warum eine Leistung besser ist als eine andere? Was haben diejenigen Studierenden, die „nur eine Drei“ bekommen haben, nicht geleistet? Sie sollten empirisch beobachtbar und für andere nachvollziehbar beschreiben können, was genau es ist, was eine Leistung „sehr gut“ oder „ausreichend“ macht. Dafür brauchen Sie Ihr [Learning Outcome](#), das die zu erwerbenden Kompetenzen genau definiert. Und Sie brauchen eine [Prüfungsaufgabe](#), die die Studierenden dazu auffordert, im Sinne des [Learning Outcomes](#) zu handeln und dadurch zu zeigen, auf welchem Niveau sie die dort beschriebenen Kompetenzen erworben haben.

Für Referate und Hausarbeiten werden häufig Notenschlüssel verwendet, die Formulierungen wie „überdurchschnittlich gut“ oder „unter den Erwartungen bleibend“ enthalten. Das wirkt objektiv, vergleicht aber lediglich die Leistungen der Studierenden in einer konkreten Gruppe miteinander. Es fehlen nachvollziehbare Beschreibungen der Leistung der Studierenden, und Sie müssen damit rechnen, dass Sie in unterschiedlichen Jahrgängen für dieselbe Leistung unterschiedliche Noten vergeben. Schließlich müssten Sie in einer sehr leistungsstarken Gruppe ja trotzdem die Note 4 vergeben, weil es immer Studierende geben wird, deren Leistungen etwas schwächer sind als die ihrer Kommiliton*innen. Wenn dann im nächsten Jahrgang die Gruppe weniger leistungsstark ist, könnte es also sein, dass jemand in dieser Gruppe herausragend ist, der in der ersten Gruppe nur Mittelmaß gewesen wäre – und dann ist Ihre Prüfungsauswertung im einen Jahr so, im nächsten anders.

Ein Beispiel

Nehmen wir an, Sie haben eine Lehrveranstaltung im Fach Werkstofftechnik angeboten, deren [Learning Outcome](#) auf der Taxonomiestufe 3 (Anwenden) liegt:

WAS Die Studierenden können Werkstoffe für spezifische, technische Anwendungen in der Technischen Gebäudeausrüstung auswählen.

WOMIT Dazu berücksichtigen Sie mechanisch-technologische Materialeigenschaften und die in einem der Gewerke Sanitär, Heizung, Klima gültigen Regelwerke, Vorschriften und Gesetze in Abhängigkeit vom Gebäudetypus (Einfamilienhaus, Mehrfamilienhaus, Hotel, Krankenhaus und Industriegebäude). Sie wenden dabei Aspekte der Wirtschaftlichkeit, Nachhaltigkeit und Ressourcenschonung an.

WOZU Dies ermöglicht ihnen die berufsspezifische, komplexe Gestaltung und Projektierung in der Technischen Gebäudeausrüstung und die nachhaltige Planung beim Thema Green Building.

Tipp

Lesen Sie auch unseren Steckbrief [Prüfungsaufgaben formulieren](#).

Die dazu passende [Prüfungsaufgabe](#) lautet, ebenfalls auf Taxonomiestufe 3:

„Sie planen die Kalt- und Warmwasserinstallation für ein Krankenhaus in Dinslaken. Ein Strangschema und Wasseranalysen liegen Ihnen vor.

1. Welche Werkstoffe sind nach DIN 1988-200 und TrinkwV für
 - a) Kaltwasser (KW)
 - b) Warmwasser (WW) zugelassen?
2. Was ist bei der Werkstoffauswahl durch begleitende Regelwerke (DVGW, VDI-Richtlinien) zu beachten?
3. Berücksichtigen Sie bei Ihrer Werkstoffauswahl die gegebene Trinkwasseranalyse, Aspekte der Hygiene und regionale Besonderheiten.

Erstellen Sie für Ihre ausgewählten Werkstoffe KW / WW (je mind. 3) ein Ranking im Bezug auf die Nachhaltig- und Wirtschaftlichkeit. Begründen Sie Ihre darauf basierende Werkstoffauswahl.“

Werkzeuge identifizieren

Für Ihr Niveaustufenmodell stellen Sie zusammen, welche „Werkzeuge“ die Studierenden anwenden müssen, um diese Aufgabe zu lösen. Nur, wenn alle diese Werkzeuge in ihrem Zusammenspiel genutzt werden, lässt sich die Aufgabe lösen – das heißt, dass Fehler beim Anwenden einzelner Aspekte Auswirkungen auf die anderen Schritte haben werden. Je taxonomisch anspruchsvoller Ihre Lehrveranstaltung ist, desto verflochtener greifen die einzelnen Werkzeuge ineinander; dies ist ein Merkmal von Tiefenlernen.

In unserem Beispiel werden berücksichtigt:

- die Regelwerke für die Auswahl der Werkstoffe
- Materialeigenschaften wie Korrosionsbeständigkeit, Materialverträglichkeit, Verbindungstechnik, Temperaturbeständigkeit
- der Gebäudetyp
- Anforderungen an Warmwasser-Installationen
- Wasserqualität und lokale Besonderheiten
- Nachhaltigkeit (z.B. Recyclingquote, Rohstoffgewinnung) und Wirtschaftlichkeit

„Damit (... Ihre...) Instrumente der deskriptiven Wahrnehmung dienen können, dürfen die Formulierungen in den skalierten Ausprägungsgraden keine analytischen Werturteile enthalten, die die eigentlichen Kriterien wieder verbergen und sofort in die Bewertung einsteigen: gut, sehr gut, angemessen, zufriedenstellend u.a. Gebrauchte Formulierungen, die in den Elementen des Learning-Outcomes Formen zeigen, wie das Learning-Outcome konkret aussehen könnte. Mit diesen Instrumenten wird der Beobachtung ein starker Filter vorgeschrieben, da die studentischen Leistungen mit diesen Instrumenten und ihren Kriterien untersucht werden.“

Maximal-, Mindest- und Regelstandard festlegen

Im nächsten Schritt beschreiben Sie die möglichen Leistungen der Studierenden in drei verschiedenen Ausprägungen: Maximal-, Mindest- und Regelstandard. Überlegen Sie für den Maximalstandard, wie eine studentische Leistung zur jeweiligen Prüfungsaufgabe aussieht, die auf ideale Weise alles zeigt, was Sie für professionell und fachlich einwandfrei halten. Den Mindeststandard beschreiben Sie, indem Sie deutlich machen, was Studierende geleistet haben, die gerade noch genug können, um zu bestehen. Der Regelstandard stellt dar, welche Bearbeitung im Durchschnitt zu erwarten sein muss.

DER MAXIMALSTANDARD beschreibt, wie eine ideale und zugleich realistisch erreichbare Aufgabenbearbeitung aussieht. Sie bewerten ja Studierende, die noch lernen und an deren Leistungen daher Maßstäbe angelegt werden müssen, die ihrem Stand im Studienverlauf entsprechen. Der Maximalstandard entspricht der Note „sehr gut“; Leistungen auf diesem Niveau werden in den Lernräumen Ihrer Veranstaltung geübt und sind damit für alle Ihre Studierenden transparent. Die manchmal vorkommende Praxis, die Note „sehr gut“ nur für Leistungen zu vergeben, die außerhalb des in Ihrer Veranstaltung Erreichbaren liegt, ist intransparent und unfair. In unserem Beispiel würde der Maximalstandard wie folgt lauten:

Der Prüfling...

- nennt die für die Werkstoffwahl relevanten Inhalte der DIN 1988-200 und wendet sie im Zusammenhang mit den zugelassenen Werkstoffen richtig an,
- berücksichtigt bei der Problemlösung begleitende Regelwerke wie z.B. DVGW-Zulassung, VDI-Richtlinien,

- kann alle Werkstoffe der TW-Installation benennen und die mögliche Werkstoffe nach der Anwendung TW kalt / warm differenzieren, entwirft Szenarien für die zusätzliche Differenzierung bei der Werkstoffauswahl, wie z.B. regionale Besonderheiten bei Einsatz bestimmter Materialien,
- lässt in seine Entscheidungsfindung die Wasserqualität und Aspekte der Hygiene einfließen,
- findet Ausschlussgründe für bestimmte Werkstoffe im Rückgriff auf seine Kenntnisse über Materialeigenschaften wie z.B. Temperaturbeständigkeit, Korrosionsbeständigkeit, Verbindungstechniken,
- entwirft Szenarien für zusätzliche Differenzierungen bei der Werkstoffauswahl, wie z.B. regionale Besonderheiten bei Einsatz bestimmter Materialien,
- wägt bei seiner Werkstoffauswahl die Vor- und Nachteile von möglichen Werkstoffen in Bezug auf die Nachhaltigkeit und Wirtschaftlichkeit ab und begründet seine darauf basierende Entscheidung der Werkstoffauswahl, zeigt in allen diesen Schritten ein systematisches Vorgehen.

DER MINDESTSTANDARD beschreibt – um bei unserem Beispiel auf Taxonomiestufe 3 zu bleiben – wie fehlerhaft eine Leistung aus dem Bereich „Anwenden“ gerade noch sein darf, um die Prüfung zu bestehen. Was ist aus fachlicher und wissenschaftlicher Sicht minimal notwendig? Mit welcher Leistung können Sie es vertreten, einem Studenten oder einer Studentin die für das weitere Studieren und den zukünftigen Eintritt in eine Fach-Community notwendige Minimalleistung zu bescheinigen, ohne die Taxonomiestufe Ihres [Learning Outcomes](#) zu unterlaufen?

Könnte der Mindeststandard erreicht werden, indem die Studierenden nicht mehr anwenden, sondern lediglich auf Stufe 2 „Verstehen“ Informationen wiedergeben, dann würde das Prüfungsergebnis eine Verzerrung mit sich bringen: die Note 4 bedeutet ja übersetzt „Die Leistung ist gerade noch im Sinne des [Learning Outcomes](#) ausreichend“ – sie muss also eine Anwendung sein, sonst ist das [Learning Outcome](#) nicht erreicht und die Prüfung darf nicht als bestanden gelten.

Der Prüfling...

- kann die für die Werkstoffwahl relevanten Inhalte der DIN 1988-200 ansatzweise erläutern und einen Zusammenhang zu den zugelassenen Werkstoffen herstellen,
- wählt Werkstoffe aufgrund von praktischen Erfahrungen, jedoch ohne detaillierte Begründung aus,
- differenziert in der technischen Anwendung TW kalt / warm nicht,
- berücksichtigt weder Wasserqualität noch Hygiene,
- berücksichtigt weder Nachhaltigkeit und Wirtschaftlichkeit.

DEN REGELSTANDARD, also eine durchschnittliche zufriedenstellende Leistung, liegt zwischen den bisher definierten Standards.

Der Prüfling...

- kann die für die Werkstoffwahl relevanten Inhalte der DIN 1988-200 erläutern und einen Zusammenhang zu den zugelassenen Werkstoffen herstellen,
- wählt Werkstoffe aufgrund von praktischen Erfahrungen, jedoch ohne detaillierte Begründung,
- differenziert in der technischen Anwendung TW kalt / warm
- berücksichtigt Wasserqualität/Hygiene oder Nachhaltigkeit/Wirtschaftlichkeit.

Wenn Sie mit diesen drei Niveaustufen (Maximal-, Minimal- und Regelstandard) noch nicht die Spannweite der studentischen Leistungen beschreiben können, ergeben sich weitere Zwischenstufen. Die Unterschiede zwischen ihnen differenzieren sich damit immer feiner aus. So entstehen – immer bezogen auf eine Taxonomiestufe – 5 bis 6 Niveaustufen.

Wenn [Learning Outcomes](#) ‚lupenrein‘ formuliert wurden, ergeben sich daraus passende [Prüfungsaufgaben](#). Um Ihre Bewertungskriterien für Ihre Studierenden genauso wie für Ihre Kolleginnen und Kollegen im Studiengang als „Abnehmer“ Ihrer Studierenden im nächsten Semester transparent zu machen, brauchen Sie ein strukturiertes Vorgehen bei der Entwicklung des Niveaustufenmodells mit einer nachvollziehbaren Definition der Werkzeuge und der Niveaustufen. Kompetenzorientierung verlangt eine messbare und an Kriterien orientierte Bewertung. Die Arbeit lohnt sich!

Literatur

1 **Reis, O. (2014):** Nicht veröff. Weiterbildungsmaterial im Rahmen der Multiplikatorenweiterbildung „Kompetenzorientierte Prüfungskonstruktionen entwickeln“ an der TH Köln 2014-2015.