

Tobias Zimmermann

Leistungsbeurteilungen an Hochschulen lernförderlich gestalten

Prüfen, Beurteilen und Rückmelden
von Lernleistungen



Verlag Barbara Budrich

Tobias Zimmermann
Leistungsbeurteilungen an Hochschulen
lernförderlich gestalten

Tobias Zimmermann

Leistungsbeurteilungen an Hochschulen lernförderlich gestalten

Prüfen, Beurteilen und Rückmelden
von Lernleistungen

Verlag Barbara Budrich
Opladen • Berlin • Toronto 2024

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über
<https://portal.dnb.de> abrufbar.

© 2024 Dieses Werk ist bei der Verlag Barbara Budrich GmbH erschienen und steht
unter der Creative Commons Lizenz Attribution 4.0 International
(CC BY 4.0): <https://creativecommons.org/licenses/by/4.0/>
Diese Lizenz erlaubt die Verbreitung, Speicherung, Vervielfältigung und Bearbeitung
unter Angabe der Urheber*innen, Rechte, Änderungen und verwendeten Lizenz.
www.budrich.de



Die Verwendung von Materialien Dritter in diesem Buch bedeutet nicht, dass diese
ebenfalls der genannten Creative-Commons-Lizenz unterliegen. Steht das verwendete
Material nicht unter der genannten Creative-Commons-Lizenz und ist die betreffende
Handlung gesetzlich nicht gestattet, ist die Einwilligung des jeweiligen
Rechteinhabers für die Weiterverwendung einzuholen. In dem vorliegenden Werk
verwendete Marken, Unternehmensnamen, allgemein beschreibende Bezeichnungen
etc. dürfen nicht frei genutzt werden. Die Rechte des jeweiligen Rechteinhabers
müssen beachtet werden, und die Nutzung unterliegt den Regeln des Markenrechts,
auch ohne gesonderten Hinweis.

Dieses Buch steht im Open-Access-Bereich der Verlagsseite zum kostenlosen
Download bereit (<https://doi.org/10.3224/84743045>).
Eine kostenpflichtige Druckversion (Print on Demand) kann über den Verlag bezogen
werden. Die Seitenzahlen in der Druck- und Onlineversion sind identisch.

ISBN 978-3-8474-3045-2 (Paperback)
eISBN 978-3-8474-1979-2 (PDF)
DOI 10.3224/84743045

Umschlaggestaltung: Bettina Lehfeldt, Kleinmachnow – www.lehfeldtgraphic.de
Titelbildnachweis: KI-generiert von DALL-E /OpenAI
Satz: Angelika Schulz, Zülpich
Lektorat: Dr. Daniela Gastell, Berlin

„Lernen ergibt sich aus dem, was die Studierenden tun und denken, und nur aus dem, was die Studierenden tun und denken. Lehrende können das Lernen nur fördern, indem sie Einfluss darauf nehmen, was die Studierenden tun, um zu lernen.“

Herbert A. Simon (zitiert nach Ambrose et al. 2010, 1, Übers. tz)

„Wettbewerbe und Prüfungen sind heutzutage in unterschiedlichsten Formaten in fast allen Lebensbereichen institutionalisiert. Sie zielen darauf ab, individuelle Leistungsunterschiede vermeintlich objektiv zu ermitteln und zu belohnen. Und genau damit stellen sie solche Unterschiede maßgeblich her.“

Nina Verheyen (2014, 67)

Dank

Das vorliegende Buch habe ich im Sinne des Urheberrechts allein verfasst. Dennoch hätte es ohne die Unterstützung durch viele Menschen nicht entstehen können. All diesen Personen gebührt mein herzlicher Dank.

Ein durch meine Arbeitgeberin, die Pädagogische Hochschule Zürich, gewährtes Weiterbildungs- und Forschungssemester verschaffte mir die nötige Zeit, um das Buch zu schreiben. Ebenso danke ich dem Verlag Barbara Budrich für die Möglichkeit, dieses Buch zu veröffentlichen, und dem institutionellen Open-Access-Fonds der PH Zürich (iOAF) für die Übernahme der Kosten für die Open-Access-Publikation.

Besonders danken möchte ich:

Gerit Thomann, der mir über Jahre hinweg die Auseinandersetzung mit dem Themenfeld der Leistungsbeurteilung ermöglicht hat und dessen Gedanken zum Thema für mich wegweisend waren. Ich danke ihm zudem für die Durchsicht des gesamten Buchmanuskripts und die daraus resultierenden wertvollen Anregungen.

Markus Weil, der mir als Vorgesetzter während meiner schreibbedingten Abwesenheit den Rücken freigehalten und das Erscheinen dieses Buches auf verschiedenen Ebenen unterstützt hat.

Meinen Kolleginnen und Kollegen aus dem Zentrum für Hochschuldidaktik und -entwicklung der Pädagogischen Hochschule Zürich: Dagmar Engfer, Mónica Feixas, Gabriel Flepp, Simone Heller-Andrist, Petra Weiss und Franziska Zellweger Moser. Dem regen fachlichen Austausch mit ihnen verdanke ich viele Einsichten, die dieses Buch geprägt haben. Zudem haben sie mir wertvolle Anregungen zum Kapitel über Grundlagen der Beurteilung gegeben, Gabriel Flepp außerdem auch zum KI-Kapitel.

Wolfgang Bühner, der große Teile des Buchmanuskripts aus der Sicht eines Hochschuldozenten kommentiert und somit geholfen hat, das Buch an den Bedürfnissen von Lehrenden aus verschiedensten wissenschaftlichen Disziplinen auszurichten.

Viviane Zimmermann für die präzise Lektüre und kritische Diskussion mehrerer Buchkapitel sowie die Unterstützung bei sozialwissenschaftlichen und statistischen Fragen.

Urs Siegfried für die anregenden Kommentare und Diskussionen zum Kapitel über Leistung und Gesellschaft sowie für unsere Gespräche über Gott und die Welt.

Susan D. Blum, Gavin T.L. Brown, Sean Kearney und Philippe Wampfler, mit denen ich parallel zum Verfassen dieses Buches im Rahmen einer Delphi-Studie inspirierende Interviews führen durfte zur Frage, wie sich

künstliche Intelligenz auf die Leistungsbeurteilung an Hochschulen auswirken wird.

Daniela Gastell für das kompetente Lektorat und Korrektorat des gesamten Manuskripts und Paula Schmieding vom Verlag Barbara Budrich für die zuverlässige Betreuung des Buchprojekts von der ersten Anfrage bis zur Veröffentlichung.

Zudem danke ich meiner Frau Viviane, unseren Kindern Nina und Jonas und meinen Freunden herzlich für ihre Unterstützung, ihre Geduld angesichts meines großen Zeitaufwands und ihre Offenheit gegenüber meiner Begeisterung für das Buchthema und das Schreiben.

Tobias Zimmermann

Inhaltsverzeichnis

Dank	7
1 Einleitung: Leistungsbeurteilung und Leistungsrückmeldung	13
1.1 Herausforderungen der Leistungsbeurteilung	13
1.2 Die Beurteilung und ihre Reflexion	18
Teil I: Grundlagen	23
2 Leistung in Gesellschaft und Bildungswesen	25
2.1 Leistung als Begriff und Wahrnehmungskonzept.....	25
2.2 Leistung und Meritokratie	30
2.3 Noten machen Leistungen	37
2.4 Metrische Fixierung und die Zertifikatsgesellschaft.....	45
2.5 Soziale Konstruktion von Lernleistung und ihrer Beurteilung	53
3 Funktionen von Leistungsbeurteilung	61
3.1 Gesellschaftliche Funktionen von Leistungsbeurteilung – Berechtigung.....	62
3.2 Didaktisch-lernbezogene Funktionen von Leistungsbeurteilung – Fördern.....	64
3.3 Dilemmata bei der Leistungsbeurteilung	67
4 Was ist Lernen und wie funktioniert es?	79
4.1 Zwei Lernexperimente.....	79
4.2 Lernen ist Konstruktion in einem Netzwerk	83
4.3 Neuronale Netzwerke	87
4.4 Drei zentrale Lernmechanismen in neuronalen Netzwerken	90
4.5 Gefühle, Motivation und Bedeutsamkeit	95
4.6 Aufbau komplexer Wissensstrukturen: Zeit und Expertise	101
4.7 Argumentatives Schlussfolgern als sozialer Prozess	109
4.8 Überleitung zum zweiten Teil des Buches.....	113

Teil II: Anwendung	115
5 Was sollen Studierende eigentlich können? Lernziele im Constructive Alignment	117
5.1 Lehre vom Ziel her denken: Constructive Alignment	118
5.2 Planungsvorlage für Lernziele	125
5.3 Übergeordnete Lernziele identifizieren und formulieren.....	127
5.4 Spezifische Lernziele identifizieren und formulieren.....	138
5.5 Epilog: Von Lernzielen und Kompetenzen.....	151
6 Voraussetzungen einer hochwertigen Leistungsbeurteilung	155
6.1 Wissen und Können als latente Konstrukte	155
6.2 Bezugsnormen der Beurteilung	158
6.3 Gütekriterien der Leistungsbeurteilung	162
6.4 Kognitive Verzerrungen und Beurteilungsfehler	170
6.5 Qualität der Leistungsbeurteilung sicherstellen.....	176
7 Formate von Leistungsnachweisen	181
7.1 Grundeigenschaften von Leistungsnachweisen	181
7.2 Aufgabenstellungen	187
7.3 Aufgabenkombinationen und Sammelnachweise	196
7.4 Strukturierungsansätze.....	199
7.5 Beschreibungsschema für Leistungsnachweisformate.....	202
7.6 Beispiele für „klassische“ Formate	204
7.7 Beispiele für innovativere Formate.....	218
7.8 Auflistung weiterer innovativer Formate.....	226
8 Formative Leistungsrückmeldung	229
8.1 Was ist eine formative Leistungsrückmeldung?	229
8.2 Leistungsrückmeldungen und Motivation	232
8.3 Inhaltliche Eigenschaften von Leistungsrückmeldungen.....	239
8.4 Feedbackinstanzen.....	246
8.5 Medialität von Feedback.....	258
8.6 Feedbackbasiertes Handeln: Schließen des Kreislaufs	265

9	Summativ beurteilen: Bewerten und Benoten	271
9.1	Die Wirkung von Bewertungen auf Motivation und Selbstregulation	272
9.2	Verschiedene Konzeptionen von Leistungsbeurteilung	274
9.3	Beurteilungsraster und andere Beurteilungsinstrumente	281
9.4	„Berechnung“ von Noten.....	298
9.5	Weniger ist mehr: Basisbewertungen und Prädikate	316
9.6	Grundlegendere Alternativen zur klassischen summativen Beurteilung	317
	Teil III: Spotlights.....	327
10	Leistungsnachweise im Curriculum	329
10.1	Bewertungen weisen Wert zu	329
10.2	Deklaratives und prozedurales Wissen im Studienverlauf.....	331
10.3	Vertikale und horizontale Integration	331
11	Leistungsbeurteilung und Prüfungsangst	335
11.1	Definition und Häufigkeit von Prüfungsangst	335
11.2	Die wichtigsten Ursachen von Prüfungsangst	337
11.3	Behandlungsmöglichkeiten von Prüfungsangst.....	338
11.4	Angstauslösende und angsthemmende Aspekte von Leistungsbeurteilung.....	339
12	Wissenschaftliche Arbeiten begleiten und beurteilen	343
12.1	Besonderheiten beim Begleiten wissenschaftlicher Arbeiten	343
12.2	Betreuungsverhältnis kontraktieren	344
13	Multiple-Choice-Prüfungen.....	347
13.1	Zum Begriff „Multiple Choice“.....	347
13.2	Was leisten MC-Prüfungen – und was nicht?	348
13.3	Struktur und Typen von MC-Aufgaben	349
13.4	Ungewollte Lösungshinweise vermeiden	352
13.5	Auswertung von MC-Prüfungen.....	353
13.6	MC-Tests als formative Selbsttests.....	354

14 Bildungsgeschichtliche Vorläufer von Leistung als schulischem Leitkonzept	357
14.1 Begriffsgeschichte	357
14.2 Ehre und Ehrtrieb als Leitkonzept der Jesuitenschulen	357
14.3 Meritentafeln und der Übergang zu Leistung als Leitkonzept.....	358
15 Künstliche Intelligenz und menschliche Leistung	361
15.1 Kulturelle Evolution oder der Mensch als Cyborg	363
15.2 Neue Arbeitsteilung zwischen Mensch und Maschine	364
15.3 Veränderungen im institutionell organisierten Lehren und Lernen.....	366
15.4 Veränderungen in der Beurteilungskultur.....	369
15.5 Offene ethische Fragen	370
16 Auflösung und Erläuterungen zu den Aufgaben in Kapitel 4.1.1.....	373
16.1 Auflösung Aufgabe 1b.....	373
16.2 Auflösung Aufgabe 2b.....	373
Literatur	375



QR-Code: Zugang zur Seite der zusätzlichen Materialien:
<https://doi.org/10.3224/84743045A>

1 Einleitung: Leistungsbeurteilung und Leistungsrückmeldung

Das Beurteilen von Leistungen, das Erteilen von Noten und das Formulieren von Leistungsrückmeldungen an Studierende sind anspruchsvolle Tätigkeiten. Viele Dozierende führen nicht gerne Leistungsnachweise durch und mögen die Notengebung noch weniger. Studierenden Feedback zu ihrem Lernfortschritt zu geben, stellt zwar für viele Lehrende kein Reizthema dar, aber im Alltag fehlt ihnen doch oft die Zeit für differenzierte Rückmeldungen. Die nachfolgenden Fallvignetten bieten anhand konkreter Beispiele Einblick in die wichtigsten Themen des vorliegenden Buches (Kap. 1.1). Im zweiten Teil der Einleitung werfe ich einen ersten kritischen Blick auf die Themen Leistung und Beurteilung, um darauf aufbauend die wichtigsten Ziele des Buches zu erläutern (Kap. 1.2).

1.1 Herausforderungen der Leistungsbeurteilung

Der Einstieg über eine Reihe von Fallvignetten gibt der sozialen Realität Raum, in der Beurteilungssituationen für alle Beteiligten auf vielschichtige Weise eingebettet sind. Die Vignetten dienen dabei als praxisnahe Beispiele, die typische Herausforderungen der Leistungsbeurteilung und Leistungsrückmeldung an Hochschulen illustrieren. Die Fälle sind bewusst gewählt, um die Dynamik und Mehrdeutigkeit der hochschulischen Lebens- und Erfahrungswelt an den Beginn dieses Buchs zu stellen und zur Reflexion anzuregen. Dies soll eine direkte Verbindung zu Ihrer eigenen Praxis herstellen und die Relevanz der Thematik greifbar machen.

Wenn Wissen nicht ausreicht

Im vergangenen Semester unterrichtete ich erstmals ein Aufbaumodul in Sozialstatistik (3. Semester, Bachelor in Politikwissenschaft). Mein Unterrichtsziel war es, den Studierenden nicht nur das Verständnis statistischer Konzepte zu vermitteln, sondern auch deren Anwendung auf reale Forschungsfragen zu fördern. Das Modul verlief angenehm, die Studierenden arbeiteten engagiert. Entsprechend hatte ich hohe Erwartungen an die Abschlussprüfung.

Die Prüfungsauswertung war deshalb für mich ein Schock: Die Leistungen der Studierenden lagen weit unter meinen Erwartungen. Die meisten Studierenden zeigten zwar ein ansprechendes Wissen über Begriffe, Konzepte und rechnerische Verfahren, hatten aber große Schwierigkeiten bei der Anwendung der Verfahrensweisen auf reale Fragestellungen. Rückblickend wurde mir bewusst, dass die Studierenden bereits im Laufe des Semesters einige Mühe gezeigt hatten, die Lerninhalte auf reale Forschungsfragen anzuwenden. Da sie jedoch die Übungen mehrheitlich rechtzeitig und vollständig bearbeiteten und gut vorbereitet in die Modulsitzungen kamen, hatte ich diesem Umstand keine große Beachtung geschenkt.

Das Prüfungsergebnis wirft für mich Fragen zur Wirksamkeit meines Lehransatzes auf: Macht mein Fokus auf Anwendung und kritisches Denken überhaupt Sinn? Was habe ich übersehen? Was kann ich beim nächsten Mal anpassen, damit die Studierenden besser für Anwendungsaufgaben gerüstet sind?

Felicia Graf, Assistentin für Politikwissenschaft

Dass Frau Graf aufgrund der schlechten Prüfungsergebnisse ihr ganzes Vorgehen hinterfragt, ist nachvollziehbar. Wenn wir einmal die Möglichkeit ausklammern, dass ihre Erwartungen an die Anwendungsfähigkeiten schlicht zu hoch waren, bieten die in diesem Buch erläuterten Konzepte folgende Erklärungsmöglichkeiten:

- Die beschriebenen Studierenden scheinen sich auf das Lernen von Begriffs- und Konzeptwissen zu fokussieren. Offenbar sind sie nicht gewohnt, erworbenes Wissen praktisch anzuwenden. Vermutlich haben sie nicht verstanden, dass dies in der Prüfung entscheidend sein würde. Hier zeigt sich, dass Lernziele definiert, kommuniziert und mit den Studierenden besprochen werden müssen. Nur so können diese wissen, welches Wissen und Können sie sich aneignen müssen (siehe Kap. 5). Damit verbunden ist das Definieren und Besprechen klarer Beurteilungskriterien, die auf diesen Lernzielen beruhen (siehe Kap. 9). Die Kenntnis und das Verstehen von Lernzielen und Beurteilungskriterien sind notwendige Voraussetzungen für Studierende, um sich gezielt auf eine Prüfung vorzubereiten. Dies trifft umso mehr zu, wenn die Anforderungen vom vertrauten Muster abweichen – so wie im geschilderten Fall, bei dem die Studierenden offenbar vor allem Prüfungen kennen, bei denen sie deklaratives Wissen auswendig lernen und wiedergeben müssen. Auswendiglernen ist aber bei Leistungsnachweisen, die auf tiefgreifendes Verstehen zielen, keine sinnvolle Lernstrategie (siehe Kap. 4.2).
- Die Fallbeschreibung lässt offen, ob und in welcher Weise die Studierenden Feedback zu den erwähnten Übungen erhalten haben. Da sie auch am

Semesterende noch große Mühe mit Anwendungsaufgaben hatten, war ihnen vermutlich die Bedeutung dieser Anforderung nicht bewusst. Das bedeutet, dass sie entweder zu wenig oder zu unklares Feedback zu den gelösten Übungen erhalten haben. Für den Lernprozess sind formative Leistungsrückmeldungen aber zentral (siehe Kap. 3 und 8).

- Eine weitere Erklärungsmöglichkeit besteht darin, dass die Anwendungsaufgaben der Abschlussprüfung deutlich von den Übungsaufgaben abweichen. In diesem Fall ergäbe sich eine Diskrepanz zwischen Lehr-Lern-Aktivitäten und Leistungsnachweis (Bruch im Constructive Alignment, vgl. Kap. 5.1).

Aus diesen Erklärungen ergeben sich einige konkrete Anpassungsideen: Felicia Graf könnte bei der nächsten Durchführung des Moduls zu Beginn des Semesters mehr Zeit aufwenden, um mit den Studierenden die Lernziele (siehe Kap. 5.2 bis 5.4) und die Beurteilungskriterien (siehe Kap. 9.3) zu besprechen. Zudem dürfte es sinnvoll sein, den Studierenden mehr Möglichkeiten zu geben, formative Rückmeldungen zu ihrem Lernen zu erhalten, sowie Anreize zu schaffen, dieses Feedback auch umzusetzen. Dabei sollte der Schwerpunkt auf Anwendungsaufgaben liegen, und insbesondere Peer-Feedback könnte eine zusätzliche Lernwirkung erzeugen (siehe Kap. 8).

Neues Format, neue Muster?

Thea Schleiden, Professorin für Zellbiologie, ist bekannt für ihre begeisterten Lehrveranstaltungen. Sie möchte den Studierenden ein tiefes Verständnis für die Zellbiologie vermitteln, führt inspirierende Diskussionen und bewirkt ein hohes Engagement der Studierenden. Diese schätzen, dass die erfolgreiche Forscherin ihnen auf Augenhöhe begegnet. An Klausuren hat Frau Schleiden allerdings wenig Interesse, da sie am Zusammenhang zwischen Klausuren und Lernfortschritt zweifelt. Sie lässt sie deshalb größtenteils von ihren Mitarbeitenden durchführen.

Als die Fakultät eine neue Studien- und Prüfungsordnung einführt, wird es möglich, Vorlesungen mit einer Posterkonferenz als Leistungsnachweis abzuschließen. Thea Schleiden sieht darin eine Möglichkeit, Klausuren zu vermeiden, und informiert die Studierenden in der ersten Vorlesungswoche über das neue Vorgehen. Während des Semesters freut sie sich, dass die Studierenden keine Fragen mehr zum Prüfungsstoff stellen.

Am Semesterende ist sie jedoch enttäuscht: Viele Poster weisen Ungenauigkeiten bei zentralen Konzepten der Zellbiologie auf und machen die zu präsentierenden Laborexperimente kaum verständlich. Dies steht im Widerspruch zur engagierten Mitarbeit der Studierenden während des Semesters. Spontan fragt die Professorin die Studierenden am Ende der Posterkonfe-

renz, ob sie ihre Poster gelungen fänden. Die teilweise emotionalen Reaktionen der Studierenden deuten darauf hin, dass viele nicht verstanden haben, wie die Poster zu gestalten sind und wie die Beurteilung erfolgt. Es wird auch deutlich, dass viele zum ersten Mal ein Poster erstellt haben und sich bei dessen Gestaltung unsicher fühlten.

Die Besprechung mit den Studierenden zeigt, dass diese nicht aus Faulheit oder mangelnder Motivation Poster gestaltet haben, die Thea Schleidens Erwartungen nicht entsprachen. Auch hier bieten sich verschiedene Erklärungsansätze an:

- Im Denken und Handeln der Professorin deutet sich ein Zielkonflikt an zwischen dem ökonomischen Einsatz ihrer Zeitressourcen und dem Gestalten eines Leistungsnachweis-Settings, das die Gütekriterien für Leistungsbeurteilungen in ansprechendem Ausmaß erfüllt (siehe Kapitel 6.3). Es liegt überdies nahe, dass Thea Schleiden die Posterkonferenz als neues Leistungsnachweisformat nicht sehr sorgfältig eingeführt hat, wodurch die Studierenden nicht genau wussten, was von ihnen erwartet wird (die Vielfalt sowie die Vor- und Nachteile von Leistungsnachweisformaten werden in Kap. 7 erläutert). Analog zur vorangehenden Fallvignette wäre es auch hier wichtig gewesen, mit den Studierenden Lernziele und Beurteilungskriterien zu besprechen (Kap. 5 und 9).
- Eine Einführung in das Gestalten und die Qualitätsmerkmale von Postern wäre notwendig gewesen, wenn viele Studierende dieses Format nicht kennen. Wie schon im vorangehenden Punkt zeigt sich hier die Wichtigkeit des Constructive Alignments: Lernziele, Leistungsnachweis und die auf den Nachweis vorbereitenden Lehr-Lern-Aktivitäten müssen aufeinander ausgerichtet sein und zueinander passen (siehe Kap. 5.1). Andernfalls ist es für Studierende schwierig, die an sie gerichteten Leistungserwartungen zu erfüllen.
- Da viele Studierende nicht mit dem wissenschaftlichen Poster als „Textsorte“ und somit auch der Posterkonferenz als Leistungsnachweisformat bekannt sind, scheint es besonders wichtig, vor einer summativen Bewertung (Note, Prädikat „bestanden/nicht bestanden“, siehe Kap. 3.1) Übungsgelegenheiten zu schaffen, bei denen die Studierenden formatives Feedback erhalten (siehe Kap. 3.2).

Die Analyse legt verschiedene Anpassungsideen nahe: Thea Schleiden könnte bei der nächsten Durchführung des Moduls zu Beginn des Semesters mehr Zeit aufwenden, um mit den Studierenden die Posterkonferenz als Format des Leistungsnachweises, die Lernziele (siehe Kap. 5.2 bis 5.4) und die Beurteilungskriterien (siehe Kap. 9.3) zu besprechen. Mit diesem bescheidenen Zusatzaufwand kann sie den Studierenden ermöglichen, wesentlich zielgerichte-

ter zu lernen – und gibt dem neuen Leistungsnachweisformat eine echte Chance, lernförderliche Wirkung zu entfalten.

Auch sollten die Studierenden genügend Möglichkeiten erhalten, das Erstellen von Postern zu üben bzw. formative Rückmeldungen zu Entwürfen des Endprodukts zu erhalten. Ähnliches gilt für das Präsentieren von Postern. Um den knappen Zeitressourcen der Professorin Rechnung zu tragen, empfiehlt sich zu diesem Zweck insbesondere der Einsatz von Peer-Feedback oder anderen Feedbackinstanzen (vgl. Kap. 8.4). Das Gespräch mit den Studierenden zeigte außerdem, wie wichtig es ist, dass Dozierende ihre Leistungsbeurteilungspraxis reflektieren und anpassen (Kap. 2 und 3).

Wessen Leistung beurteile ich?

Vor einiger Zeit begann ich, Bachelorarbeiten im Studiengang Physiotherapie zu betreuen. Viele Studierende möchten ihre Abschlussarbeit einfach möglichst effizient und mit einer passablen Note hinter sich bringen. Doch kürzlich traf ich auf eine außergewöhnlich motivierte Studentin. Nuriya wollte in ihrer Bachelorarbeit die Auswirkungen spezifischer motorischer Übungen auf die neuromuskuläre Koordination untersuchen. Ich warnte sie, dass ihr sehr forschungsorientiertes Vorhaben für eine Bachelorarbeit aufwändig sei. Aber sie war begeistert vom Thema und der Aufwand schreckte sie nicht ab.

Nuryias Engagement war bemerkenswert. Sie recherchierte unermüdlich, besuchte Schulungen in Forschungsmethodik und führte Tests mit selbst rekrutierten freiwilligen Probanden durch. Regelmäßig kam sie in meine Sprechstunde. Ich beantwortete ihre offenen Fragen, wies sie auf neue Publikationen hin und unterstützte sie bei der methodischen Planung der Tests.

Als sie ihre Arbeit einreichte, war ich zuerst begeistert von deren Qualität. Doch als ich das Erstgutachten zur Arbeit verfassen sollte, wurde ich unsicher. Aufgrund unseres intensiven Austausches fiel es mir schwer, die Grenze zwischen ihrem eigenen Beitrag und meinen Ratschlägen zu erkennen. Hatte ich als Dozent zu stark in den Arbeits- und Lernprozess der Studentin eingegriffen? Inwiefern konnte ich ihre Arbeit noch unabhängig beurteilen und bewerten? Und sollte ich mich künftig beim Begleiten von Bachelorarbeiten stärker abgrenzen?

Mateo Precioso, Dozent für Neuromotorik

Die Zweifel von Herrn Precioso verweisen auf grundlegende Herausforderungen im Zusammenhang mit dem Beurteilen von Studierendenleistungen:

- Zwischen dem Beraten und Begleiten von Studierenden und dem abschließenden Bewerten ergibt sich ein Rollenkonflikt: Als Berater (wie auch im Lehralltag) soll ein Dozent die Studierenden unterstützen, als Bewerter muss er Stärken und Schwächen möglichst unabhängig feststellen und mitunter harte „Urteile“ fällen (siehe Kap. 3.3 zu Rollenkonflikten).
- Die Studentin war offenbar sehr motiviert und engagiert, hat aber gleichzeitig extensiv auf die Unterstützung des Dozenten zurückgegriffen. Dies wirft Fragen danach auf, wie Studierende dazu ermutigt und befähigt werden können, ihre eigenen Lernprozesse zu reflektieren und dadurch ihre Metakognition und Selbststeuerung weiterzuentwickeln (siehe Kap. 8).

Viele Lehrende machen sich gelegentlich ähnliche Gedanken wie Mateo Precioso. Darin äußern sich Widersprüche, die Lehrsituationen strukturell inne- wohnen und sich nicht auflösen lassen, etwa der Widerspruch, dass Studierende zum selbständigen fachlichen Denken *geführt* werden sollen. So fragte der Philosoph Immanuel Kant (vgl. 1803, 32) schon vor mehr als 200 Jahren, ob Menschen durch Zwang zur Freiheit erzogen werden könnten. Das vorliegende Buch möchte Ihnen in diesem Zusammenhang helfen, Ihren eigenen, bewussten Umgang mit diesen strukturell bedingten Herausforderungen zu finden (siehe Kap. 3).

Die Erfahrung von Mateo Precioso zeigt des Weiteren die Komplexität von Betreuungsverhältnissen auf, etwa bei der Begleitung wissenschaftlicher Arbeiten oder Projekte. Es lohnt sich in solchen Fällen, die gegenseitigen Erwartungen und Verantwortlichkeiten durch eine schriftliche Vereinbarung vorab zu klären (siehe Kap. 12).

1.2 Die Beurteilung und ihre Reflexion

Welche Erlebnisse kommen Ihnen zuerst in den Sinn, wenn Sie das Wort „Prüfungen“ hören? Schulische Situationen? Lebensprüfungen, die Ihnen vom Schicksal zugemutet wurden? Die Prüfung von Bilanzen und Geschäftsunterlagen? Was verstehen Sie unter einer Leistung? Wer erbringt Ihres Erachtens eine Leistung, und wie kann man sie erfassen, bewerten, vergleichen, beurteilen?

Begriffsklärung

Im Zusammenhang mit Leistungsbewertung an Hochschulen ist oft von „Prüfungen“ die Rede. Dieser Begriff beinhaltet aber diverse Einschränkungen – bereits eine wissenschaftliche Arbeit dürfte kaum als Prüfung bezeichnet werden. Deshalb wird in diesem Buch der Begriff *Leistungsnachweis* als Oberbegriff für alle Formen von Leistungsüberprüfung verwendet.

Die Kognitionen und Emotionen, die Sie mit diesen ganz persönlichen Schlüsselerlebnissen verbinden, sowie Ihre bewussten und unbewussten Schlussfolgerungen daraus dürften Ihren Umgang mit Beurteilungssituationen wesentlich prägen. Betrachten Sie Leistungsnachweise (siehe obenstehende Begriffsklärung) als lästiges Übel, das sowohl Sie als auch Ihre Studierenden irgendwie hinter sich bringen müssen? Oder glauben Sie, dass Leistungsnachweise eine wichtige Rolle spielen, indem sie die schlechten Studierenden aussortieren helfen, so dass nur die geeigneten übrigbleiben? Oder sehen Sie sich eher in der Rolle der Beraterin, welche die Studierenden beim Lernen und beim Erfüllen der durch die Institution Hochschule gestellten Anforderungen unterstützt? Gibt es bestimmte Formen von Leistungsnachweisen oder Aufgabentypen, die Sie in Ihrem Fach für besonders (un)geeignet halten, oder empfehlen Sie Ihren Studierenden bestimmte Vorgehensweisen zur Prüfungsvorbereitung?

Auch wenn Sie diese Haltungen und Überzeugungen auf Ihre Erfahrungen als lehrende und/oder lernende Person sowie auf Wissen über Lehren, Lernen und Prüfen stützen können: Subjektiv geprägte, teilweise unbewusste Einschätzungen und die mit ihnen verbundenen Emotionen prägen unser Denken und Handeln stark – oft mehr als kognitive Wissensbestände. Das gilt selbst dann, wenn dieses Wissen über lange Zeit in einem fachlichen und/oder professionellen Kontext aufgebaut wurde und relativ weit reicht (vgl. Wahl 2001). Dies erklärt, weshalb selbst eine Professorin für Erziehungswissenschaft eine monologische Vorlesung über den Wert dialogischer Lernformen halten kann – die langjährige akademische Prägung führt zu einem unhinterfragten Bild, wie eine Vorlesung abzulaufen hat, und erzeugt eine kognitive Dissonanz zwischen vermitteltem Inhalt und Vermittlungsform.¹

- 1 Das in Kap. 5.1.2 vorgestellte Modell des Constructive Alignments ermöglicht, das geschilderte Problem didaktisch genauer zu fassen: Sollen Studierende den Wert dialogischer Lernformen erkennen, korrespondiert dieses Lernziel nicht mit der Unterrichtsmethode des frontalen Vortrags – es liegt ein Bruch im Alignment vor.

Das vorliegende Buch möchte deshalb mehr sein als nur eine „Best-Practice“-Anleitung zum effizienten und lernförderlichen Durchführen von Leistungsbeurteilungen. Es möchte Sie darüber hinaus dazu auffordern, Ihre Einstellungen und Praktiken im Umgang mit Leistungsbeurteilungen sowie beim Erteilen von Leistungsrückmeldungen an Studierende kritisch zu hinterfragen. Das beginnt schon bei der Frage danach, was im Zusammenhang mit Lernen überhaupt unter einer „Leistung“ zu verstehen ist (siehe Kap. 2). Hier ist auch die Beziehung zu den Menschen im Auge zu behalten, um deren Lernen es geht: „Lernen ergibt sich aus dem, was die Studierenden tun und denken, und nur aus dem, was die Studierenden tun und denken. Lehrende können das Lernen nur fördern, indem sie Einfluss darauf nehmen, was die Studierenden tun, um zu lernen“ (Herbert A. Simon, zitiert nach Ambrose et al. 2010, 1, Übers. tz).

Deshalb ist zu beachten, dass Leistungsnachweise nicht nur zur Bilanzierung von Lernleistungen dienen, sondern ganz wesentlichen Einfluss darauf ausüben, wie Studierende lernen (siehe Kap. 5). Damit sind wiederum Fragen nach der gesellschaftlichen Aufgabe von Hochschulen als Institutionen verbunden. Denn das Lernen an Hochschulen endet mit der Verleihung von Studienabschlüssen, welche zur Übernahme als besonders relevant eingeschätzter gesellschaftlicher Rollen berechtigen, wie Ärztinnen, Juristen, Ingenieurinnen oder Lehrer. Unsere Beurteilungspraxis bewegt sich deshalb zwischen dem Fördern von Lernen, der Sicherstellung des Erwerbs von Wissen und Können sowie dem Erteilen von gesellschaftlichen Berechtigungen (siehe Kap. 3).

Das kritische Hinterfragen Ihrer Beurteilungspraxis macht Sie im Erfolgsfall zum „Reflective Practitioner“ im Sinne von Schön (1991). Indem Sie Ihren Umgang mit Leistungsnachweisen und Leistungsrückmeldungen auf verschiedenen Ebenen reflektieren können, erweitern Sie Ihre Handlungsalternativen: Sie können beim Entwickeln von Leistungsnachweisen situationsadäquat Formate und Aufgaben auswählen und gestalten, auf der institutionellen Ebene an der Erarbeitung von lernförderlichen Regelungen in Bezug auf die Bewertung und Beurteilung von Leistungen mitwirken und Studierende wirksam bei der Vorbereitung auf Leistungskontrollen begleiten.

Bei aller kritischer Distanz zum Leistungsbegriff sowie zur Bewertung und Beurteilung von Leistungen sei somit gesagt: Wenn in diesem Buch der Leistungsbegriff hinterfragt und dysfunktionale Prüfungs- und Bewertungspraktiken (z. B. das Erteilen von Noten) moniert werden, dann immer mit dem Ziel, durch eine ganzheitlichere Praxis der Leistungsbeurteilung und -rückmeldung *an Hochschulen mehr und besseres Lernen zu ermöglichen*. Ich bin überzeugt, dass die Hochschulen dadurch auch ihre gesellschaftlichen Funktionen als Bildungs- und Forschungsstätten sowie als Qualifikationsinstanzen noch wirkungsvoller wahrnehmen.

So ist mein unmittelbares Ziel, Ihnen mit diesem Buch hilfreiche praktische Ideen für die Gestaltung von Beurteilungssituationen und den Umgang mit ihnen an die Hand zu geben. Darüber hinaus hege ich die vielleicht verwegene Hoffnung, dass es ein wenig dazu beiträgt, Ihre persönliche Beurteilungspraxis – und mittelbar vielleicht diejenige ganzer Bildungsinstitutionen – lernförderlicher, aussagekräftiger und gerechter zu machen.

Teil I: Grundlagen

2 Leistung in Gesellschaft und Bildungswesen

Sie möchten wissen, wie Sie die Leistungen Ihrer Studierenden effizienter oder zuverlässiger erfassen und beurteilen können und/oder wie Sie den Studierenden möglichst lernwirksames Feedback zu den gezeigten Lernleistungen geben können? Das Wissen, das Sie sich in diesem Kapitel aneignen können, wird einerseits vermutlich Ihre Vorstellungen von „Zuverlässigkeit“ im Bereich der Leistungsmessung und -beurteilung verändern. Andererseits ermöglicht es Ihnen ein neues Bild davon, wie studentische Leistungen zustande kommen und interpretiert werden. Dies ist eine notwendige Vorbedingung dafür, diese Leistungen angemessen erfassen, beurteilen, bewerten und Feedback dazu geben zu können.

Lassen Sie sich deshalb auf eine Reise durch die Begriffsgeschichte der Leistung sowie gesellschaftlicher Leitkonzepte im Umgang mit Leistung mitnehmen. Wir gehen zuerst der Rolle von Leistung als Wahrnehmungskonzept in unserer Gesellschaft und speziell im Bildungswesen auf den Grund (Kap. 2.1). Auf dieser Basis können wir das Leitkonzept der Meritokratie kritisch analysieren (Kap. 2.2) und die Zuverlässigkeit und Auswirkung von Noten hinterfragen (Kap. 2.3). Dadurch stoßen wir auf Phänomene, die sich durch eine ausgeprägte gesellschaftliche Orientierung an Messbarkeit und Vergleichen sowie dem Erstellen von Rangordnungen erklären lassen (Kap. 2.4). Dabei werden auch Zusammenhänge zwischen der Beurteilung von studentischen Leistungen und der Beurteilung von Hochschulen als Institutionen oder von Forschungsleistungen sichtbar. Der kritische Blick auf Leitkonzepte wie Meritokratie, Metrisierung und Kredentialismus ermöglicht es schließlich, den Leistungsbegriff einer für Lehr- und Lernzwecke konstruktiven Sichtweise zuzuführen (Kap. 2.5).

2.1 Leistung als Begriff und Wahrnehmungskonzept

Überraschenderweise zeigt die Begriffsgeschichte, dass der Leistungsbegriff im Bildungskontext noch nicht so lange verwendet wird, wie man aufgrund seiner heutigen Allgegenwärtigkeit vermuten könnte. Er erweist sich nicht nur als eng verbunden mit technologischem Fortschritt, sondern auch mit Vorstellungen von Mess- und Vergleichbarkeit. „Leistung“ ist heute zwar ein vieldeutiger Begriff. Diese Vieldeutigkeit ist aber keineswegs beliebig, sondern führt ins Zentrum unseres gesellschaftlichen Selbstverständnisses. „Leistung“ kann als ein Wahrnehmungskonzept beschrieben werden, das

wesentliche Vorgänge in Gesellschaft und Bildungswesen steuert (Reh, Berdelmann u. Scholz 2015, 37).

Heute umfasst der Begriff „Leistung“ im allgemeinen Sprachgebrauch gewöhnlich vier Bedeutungselemente, die je nach Verwendungszusammenhang mehr oder weniger im Vordergrund stehen (vgl. Schlie 1988, 63f.):

- Leisten ist zielgerichtetes Handeln, wobei das Ziel durch die Handelnden oder von außen gesetzt sein kann.
- Damit verbunden ist eine gewisse Anstrengung, eine Leistung stellt sich nicht einfach von selbst ein.
- Als Ergebnis des Leistens wird ein bestimmtes Resultat erwartet, zu dem die zielgerichtete Anstrengung führen soll.
- Das Ergebnis wird in der Regel beurteilt, meistens anhand der gesetzten Ziele, auf die sich das Leisten bezieht, und manchmal auch aufgrund des Grades der investierten Anstrengung. Leistung kann folglich anhand von Gütekriterien beurteilt werden.

Diese Elemente treten auch – bis in die begriffliche Ebene – in der heutigen Konzeption von Schulleistung hervor, in der von Lernzielen, learning outcomes und Gütekriterien der Beurteilung die Rede ist (siehe Kap. 5 und 6). Das ist kein Zufall und zeigt: Der heutige Leistungsbegriff ist auch für den Bildungsbereich prägend. Das gesellschaftliche und das bildungsbezogene Leistungsverständnis sind sogar untrennbar miteinander verwoben, wie die folgenden Seiten deutlich machen.

Dass Leistung nicht immer so stark mit Fragen der Bildung gekoppelt war, zeigt ein Blick in die Geschichte des Schulwesens, den Kapitel 14 im Rahmen eines bildungshistorischen Exkurses vertieft. Erst im Laufe des 19. Jahrhunderts rückte die auf das fachliche Lernen bezogene „Leistung“ in den Fokus. So heißt es in der „Enzyklopädie des gesamten Erziehungs- und Unterrichtswesens“ 1887, die Zensur (=Note) diene dazu, die Schüler „zu einem richtigen Urteil über sich selbst, [ihre] Fähigkeiten und Leistungen anzuleiten“ (zit. nach Ziegenspeck 1999, 75).² Im Rahmen dieser Entwicklung wurden schulische Leistungen im Deutschland des 19. Jahrhunderts zunehmend in Zeugnisnoten abgebildet, wobei die Zahl der Abstufungen noch variierte, um sich dann im 20. Jahrhundert zu stabilisieren (vergleichbar mit den Entwicklungen in den USA, vgl. Brookhart et al. 2016, 805).³

- 2 Als eigenständiges Stichwort tritt „Leistung“ allerdings erst seit 1930 in pädagogischen Handbüchern und Enzyklopädien auf (vgl. Furck 1972, 140, FN 34).
- 3 Die heute in Deutschland noch gültige sechsstufige Skala von 1 (=sehr gut) über 4 (=ausreichend) bis 6 (=ungenügend) wurde schließlich 1938 in der Zeit des Nationalsozialismus eingeführt und erstaunlicherweise seither in der Bundesrepublik Deutschland beibehalten. In Österreich wird dagegen heute eine Skala

Die Entstehung von Schulnoten im Zeitraum von 1800 bis ca. 1950 ist Symptom eines grundsätzlichen Wandels im Bildungswesen. Erst durch die Einführung und nationale Standardisierung von Schulnoten wurde eine Art Währung eingeführt, mit der Schulleistungen scheinbar unabhängig vom Leistungsniveau einer Klasse oder Schule vergleichbar wurden (vgl. Verheyen 2018, 63).

Interessanterweise verlief die Entwicklung des modernen Leistungssports parallel zur Standardisierung der Schulnoten. Auch im Sport wurden Leistungen zwar schon früher verglichen (vgl. Mallinckrodt u. Schattner 2016). Aber erst ab der zweiten Hälfte des 19. Jahrhunderts wurde zunehmend situationsübergreifend gemessen und verglichen. „Damit schien es möglich, einen Wettbewerb überall auf der Welt und beliebig häufig zu wiederholen und *alle* erreichten Ergebnisse von *allen* Sportlern – Frauen waren mal wieder größtenteils ausgeschlossen – miteinander zu vergleichen, ganz egal, wo und wann sie erbracht worden waren, ob die Sonne schien oder den Athleten Regen ins Gesicht peitschte. (...) Das ging bis hinauf zum ‚Weltrekord‘ (...)“ (Verheyen 2018, 87, Hervorhebungen von Verheyen).

Hier zeigt sich eine klare Gemeinsamkeit von sportlicher und schulischer Leistungsmessung: Beide zielen auf eine globale Vergleichbarkeit der gemessenen Leistungen (vgl. Heintz 2016). Diesen Megatrend erkannte schon der Philosoph Friedrich Nietzsche, als er 1878 die Moderne zum „Zeitalter der Vergleichung“ ausrief (Nietzsche 1999 [1878], 45). Und der Trend beschränkte sich mitnichten auf Schule und Sport, sondern war eng verbunden mit technischen, wirtschaftlichen und politischen Entwicklungen. Dank gesellschaftlicher und technologischer Entwicklungen waren nicht nur immer genauere Messungen von physikalischen und anderen Werten möglich, sondern diese konnten durch zunehmend leistungsfähigere Kommunikationstechnologien immer schneller und globaler übermittelt und verglichen werden.⁴ Solche Prozesse ermöglichten auch die Entstehung von Börsenindizes

von 1 (=sehr gut) über 4 (=genügend) bis 5 (=nicht genügend) verwendet (vgl. Ziegenspeck 1999, 76–79 u. 95). Die Deutschschweiz benotet hingegen in umgekehrter Richtung und mit anderer Gewichtung von 6 (=sehr gut) über 4 (=genügend) bis zu 1 (=schlecht). Man beachte, dass der Anteil der ungenügenden Noten am gesamten Notenband von 25 % im österreichischen bis zu 60 % im Deutschschweizer System variiert. Dies führt z. B. in der Schweiz dazu, dass der ungenügende Teil der Bewertungsskala kaum genutzt wird – Noten unter 3 sind sehr selten und zeigen außerordentlich schwache Leistungen an.

- 4 Die Zusammenhänge von Ursache und Wirkung dieser bis heute anhaltenden Dynamik werden schnell komplex und unübersichtlich: Haben gesellschaftliche Entwicklungen erst den rasanten technologischen Fortschritt ab dem 19. Jh. ermöglicht, etwa die Einführung des Meters als Maßeinheit in der Folge der Französischen Revolution (vgl. Konersmann 2021, 205 und 298–300 (FN 117))? Oder waren die technologischen Fortschritte, im 19. Jh. z. B. industriell produ-

wie dem Dow Jones oder die Abwicklung von Außenhandelsgeschäften ohne persönliche Vertrauensbeziehungen (vgl. Gill 2017, 393; Ambrosius 2018, 285).

Gerade auch aufgrund der Wirtschaftsnähe des Leistungsdenkens wird oft postuliert, dass es aus der Verbürgerlichung der Gesellschaft hervorging, als „bürgerliches Leistungsdenken“. Die Zusammenhänge sind freilich wesentlich komplexer, zumal die Entwicklungen in verschiedenen Gesellschaften auf unterschiedlichen Kontinenten fast zeitgleich stattfanden (vgl. für nähere Erläuterungen Verheyen 2018, 2014). Sapolsky (2018, 296f.) weist darauf hin, dass die im 19. Jahrhundert einsetzende massive Urbanisierung, die von 1870–1914 einen ersten Höhepunkt erreichte, anthropologisch weitreichende Folgen hatte.

Laut dem Anthropologen Sapolsky ist es unter Primaten erstmalig, dass sich regelmäßig Fremde begegnen, die sich nie wieder sehen – eine Großstadtbewohnerin begegnet täglich wesentlich mehr Fremden als Bekannten. Dies habe zur Erfindung der anonymen Handlung geführt. Henrich et al. (2010) weisen in diesem Zusammenhang darauf hin, wie erstaunlich es ist, dass in anonymen Massengesellschaften Fremde mit hoher Zuverlässigkeit miteinander Transaktionen zum gegenseitigen Vorteil durchführen.⁵ Die Entwicklung von übergreifenden Normen und Institutionen dürfte nötig geworden sein, um Fairness in solchen flüchtigen Austauschbeziehungen zu gewährleisten. Ein Aspekt davon betrifft die zunehmende Organisation gesellschaftlichen Austausches als Markt: In der Tat geht laut Henrich et al. (2010) eine größere Vermarktlichung von Gesellschaften mit mehr Fairness im Verhalten bei an Marktmechanismen orientierten Spielen einher. Märkte wiederum leben vom Vergleich und legen damit Quantifizierungen nahe: Die zum Vergleich „synonyme Formulierung ‚aneinander messen‘ stellt klar, worum es geht. Dinge vergleichen heißt, sie unter dem Aspekt ihrer Gleichartigkeit zu betrachten, und das bedeutet: sie schon mit der Art des Zugriffs vergleichbar zu machen“ (Konersmann 2021, 52). Die Leistung als quantifiziertes Maß menschlicher Tätigkeit wurde somit zur Basis vieler gesellschaftlicher Bezie-

zierte Uhren und die Telegraphie, verantwortlich für gesellschaftliche Entwicklungen wie die Urbanisierung? Diese beiden Aspekte scheinen untrennbar miteinander verwoben.

- 5 Auch die Erfindung der Kriminalliteratur, die in die gleiche Zeit fällt, sieht Sapolsky (2018, 296) als Auswirkung der neuartigen Urbanisierung: Erst das Setting von anonymen, städtisch geprägten Gesellschaften macht die Frage spannend, wer ein bestimmtes Verbrechen begangen hat – in auf persönlichen Kontakten basierenden Gesellschaftsordnungen in weniger dicht besiedelten Gegenden wissen in der Regel alle gegenseitig über ihre Handlungen Bescheid.

hungen – mit weitreichenden Folgen, von denen wir in diesem Kapitel nur die Spitze des Eisbergs erkunden können.⁶

Zugleich ist es wichtig zu beachten, dass viele der entstandenen Maße für die Messung und den Vergleich von Leistungen keine wirklich objektiven und absoluten Vergleiche ermöglichen – die Genauigkeit der Maße täuscht über die Unterschiedlichkeit der Bedingungen hinweg, unter denen die Leistungen entstanden sind. Um nochmals zum Sport zurückzukehren: Es spielt für die Gültigkeit eines Marathonweltrekords keine Rolle, welche Temperatur- oder Windverhältnisse bei seiner Erbringung geherrscht haben. Und innerhalb eines Radzeitfahrens oder Skirennens können ganz unterschiedliche Wind-, Niederschlags- und Lichtverhältnisse herrschen – die Zeiten werden dennoch auf Hundertstelsekunden genau gemessen und verglichen. Falls Sie nun einwenden möchten, dass im Sport auch die Show im Vordergrund steht und die Leistungsmessung und -beurteilung im Bildungsbereich gewiss seriöser erfolgt: Leider muss ich Sie enttäuschen. Die aktuelle Praxis der Beurteilung und des Vergleichs von Lernleistungen – nicht nur, aber ganz besonders durch Noten – ist noch viel unzuverlässiger. Die widersprüchlichen Zusammenhänge von Messung, Beleg und Vergleich von Leistungen, die sich in Gesellschaft und Bildungswesen etabliert haben, zeige ich nachfolgend auf. Am Ende des Kapitels wird auf dieser Basis dargelegt, inwiefern Lernleistungen durch Bildungseinrichtungen wie Hochschulen (auch) sozial konstruiert werden. Dadurch lässt sich eine differenzierte Definition des Begriffs „Lernleistung“ vornehmen (siehe Kap. 2.5.2).

- 6 Sapolsky weist angesichts der Studie von Henrich, Joseph et al. (2010) darauf hin, dass wir dazu neigen, Marktmechanismen als den Gipfel der Regulation von komplexen menschlichen Bedürfnissen zu betrachten – alles kann zu einem Geldwert abstrahiert werden. Doch anthropologisch betrachtet sei menschliche Gegenseitigkeit ursprünglich wesentlich intuitiver und ganzheitlicher: „Der Typ hier ist ein großartiger Jäger; der andere dort kann das weniger gut, aber hält dir den Rücken frei, wenn ein Löwe in der Nähe ist; jene Frau dort ist enorm gut darin, die besten Mongongo-Nüsse zu finden, während diese ältere Frau alles über Heilkräuter weiß, und die seltsame Type dort drüben kann die besten Geschichten erzählen. Wir wissen, wo die anderen leben, mit der Zeit gleichen sich die Schuldigkeiten aus, und wenn jemand wirklich das System missbraucht, werden wir uns gemeinsam darum kümmern“ (Sapolsky 2018, 500, Übers. tz). Deshalb sieht er Marktinteraktionen als Verarmung der Gegenseitigkeit menschlichen Handelns.