

Adriane Langela-Bickenbach
Ricarda Dreier
Philippe Wampfler
Christian Albrecht (Hrsg.)

PÄDAGOGIK

Wege zu einer zeitgemäßen Prüfungskultur

Grundlagen und Praxisbeispiele



E-Book inside

BELTZ

Adriane Langela-Bickenbach ist Gymnasiallehrerin für Fremdsprachen und Fachleiterin Englisch in der Lehrerausbildung am ZfsL Aachen.

Ricarda Dreier ist Gymnasiallehrerin für Deutsch und Latein sowie Fachberaterin »Bildung in der digitalen Welt« in der Bezirksregierung Detmold.

Philippe Wampfler ist Deutschlehrer an der Kantonsschule Uetikon am See (Schweiz) und Dozent für Deutschdidaktik an der Universität Zürich.

Dr. Christian Albrecht ist wissenschaftlicher Mitarbeiter am Lehrstuhl für Didaktik der deutschen Sprache und Literatur an der Friedrich-Alexander-Universität Erlangen-Nürnberg.

Die Herausgeber:innen sind Mitglieder des Instituts für zeitgemäße Prüfungskultur.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt.
Jede Verwertung ist ohne Zustimmung des Verlags unzulässig.
Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen
und die Einspeicherung und Verarbeitung in elektronische Systeme.



Dieses Buch ist erhältlich als:
ISBN 978-3-407-63311-8 Print
ISBN 978-3-407-63321-7 E-Book (PDF)

1. Auflage 2024

© 2024 Beltz
in der Verlagsgruppe Beltz · Weinheim Basel
Werderstraße 10, 69469 Weinheim
Alle Rechte vorbehalten

Lektorat: Dr. Erik Zyber
Umschlaggestaltung: Michael Matl
Umschlagabbildung: getty images © RoJDesign

Herstellung und Satz: Jenny Pötzsch
Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza
Beltz Grafische Betriebe ist ein klimaneutrales Unternehmen (ID 15985-2104-100).
Printed in Germany

Weitere Informationen zu unseren Autor:innen und Titeln finden Sie unter: www.beltz.de

Inhalt

Vorwort	8
----------------------	---

Christian Albrecht

Aufgaben und Herausforderungen einer zeitgemäßen Prüfungskultur am Beispiel von 4K und KI	12
--	----

I Probleme der traditionellen Prüfungskultur

Philippe Wampfler

Meritokratie und individuelle Leistung als Märchen für Erwachsene	34
--	----

Philippe Wampfler

Eigenständigkeit, Vergleichbarkeit, Allokation	39
---	----

Lars Mecklenburg und Philippe Wampfler

Von der Schwierigkeit, mit Prüfungen zu messen	43
---	----

Lars Zumbansen

Projektbericht: Mitmachausstellung	48
---	----

Hendrik Haverkamp und Julian Schmitz

Psychische Gesundheit und Prüfungskultur	51
---	----

Hendrik Haverkamp

Projektbericht: Eine Klassenarbeit ohne Aufsicht	57
---	----

II Argumente für eine neue Prüfungskultur

Björn Nölte

Kollaboration ermöglichen und bewerten	60
---	----

Uta Hauck-Thum

Zukunftsorientierter Kompetenzerwerb beim Übertritt in die Sekundarstufe	68
---	----

Lea Schulz

Inklusion	73
------------------------	----

Anna Reuter

Leistungsräume als Lerngelegenheiten Inwiefern die aktuelle Prüfungskultur Reflexion als Lernpotenzial missachtet	78
--	----

Christian Stalder

Entwicklungsorientiert prüfen	86
--	----

Björn Nölte

Feedback und formatives Assessment	91
---	----

Anne Sliwka/Janina Beigel

Projektbericht: Deeper-Learning-Einheit »Klimawandel – Klimafolgen«	98
---	----

Anne Sliwka und Janina Beigel

Warum Deeper Learning authentische Formen der Leistungsbewertung und -rückmeldung erfordert	101
--	-----

Anne Sliwka/Janina Beigel

Projektbericht: Deeper-Learning-Einheit »Mikroplastik in der Antarktis«	108
---	-----

Silvia-Iris Beutel

Demokratiepädagogik, Partizipation und neue Prüfungskultur	111
---	-----

Felix Winter

Das Dialogische Lernkonzept Prüfen und Fördern integrieren	115
--	-----

Hendrik Haverkamp und Philippe Wampfler

Die Nutzung von KI	132
---------------------------------	-----

Hendrik Haverkamp

Projektbericht: Eine Klassenarbeit mit KI	136
--	-----

III Konkrete Gestaltung einer neuen Prüfungskultur

Lars Mecklenburg

Schieberegler und Formatspinne 140

Stefanie Lotz

Projektbericht: Projektarbeit »Future Scenarios« 146

David Tepasse

Projektbericht: Spieleentwicklung mit Miro 149

Philippe Wampfler

Prüfungsformate, Mediennutzung und Arbeitsprozesse 151

Adriane Langela-Bickenbach

#GLASKonzept – virtual exchange Projekt 155

Oliver Schmitz

**Zeitgemäße Prüfungskultur aus der Perspektive
des Schulrechts** 159

Dejan Mihajlović und Philippe Wampfler

Kommunikation mit Eltern 164

Philippe Wampfler

Fortbildung im Kollegium 170

Niels Winkelmann

**Projektbericht: Lernprozesse eigenständig fortsetzen:
»Religiosität heute – ganz persönlich«** 175

Tobias Meyer

**Potenziale und Herausforderungen zeitgemäßer
Prüfungsformate im Fach Deutsch** 178

Epilog: Das Institut für zeitgemäße Prüfungskultur 183

Vorwort

Die Notwendigkeit einer zeitgemäßen Prüfungskultur

Dieses Buch dokumentiert einen Wandel der schulischen Prüfungskultur. Dieser Wandel ist real, hat viele Facetten und verschiedene Ursachen. Davon zeugen die unterschiedlichen Perspektiven und vielfältigen Beispiele, die in diesem Band versammelt sind.

Grundlegend für den Wandel ist die Einsicht, dass die klassische schulische Prüfungskultur, bei der mit einer Sammlung von Aufgaben zu einem bestimmten Zeitpunkt individuelle Leistungen abgefragt werden, nicht funktioniert – und nie funktioniert hat. Das bedeutet: Die verbreiteten Vorstellungen, dass Prüfungen Lernende motivieren, über den Vergleich ihrer Leistungen wichtige Erkenntnisse hervorzubringen, oder dass sich individuelle Leistungen überhaupt messen lassen, sind falsch. Die traditionelle Prüfungskultur geht von Annahmen aus, die einer kritischen Evaluation nicht standhalten. Sie funktioniert aber auch in einer weiteren Hinsicht nicht: Klausuren beschädigen pädagogische Beziehungen und die Lernbereitschaft sowie Eigenverantwortung von jungen Menschen. Sie führen nicht nur zu unnötiger Frustration, sondern auch zu massivem Stress und erschweren dadurch die schulische Arbeit.

Pointiert gesagt: Die Notwendigkeit einer neuen Prüfungskultur ergibt sich aus der Einsicht, dass Lernen im Mittelpunkt von Unterricht stehen muss. Sobald analysiert wird, wie Lernprozesse ablaufen und wie gute Schulen und Lehrer:innen sie unterstützen, wird deutlich, dass die etablierten Formen der Leistungsmessung die Qualität der Lernprozesse massiv beeinträchtigen.

Im Laufe der Lektüre werden Sie viele weitere Gründe finden, die einen Wandel der Prüfungskultur notwendig machen: etwa die Arbeitsformen in einer Kultur der Digitalität, die alle Voraussetzungen klassischer Leistungsmessung obsolet werden lassen. In einer von Vernetzung und Gemeinschaftlichkeit geprägten Lebenswelt erbringt beispielsweise niemand Leistungen isoliert und ohne Zugriff auf Informationen in verschiedenen Medien. Oder die psychische Gesundheit junger Menschen, für die Vertrauen, Selbstwirksamkeit und Resilienz bedeutsam, Stress und Beurteilungen aber schädlich sind.

Trotzdem werden Berührungsängste mit einer neuen Prüfungskultur bestehen bleiben. Auch das hat einen Grund. Schule ist tief in unsere Sozialisierung einge-

bettet. Wir haben alle erfahren, dass es wichtig ist, bei Prüfungen Leistungen zu erbringen. Wir haben versagt, gelitten, aber auch Erfolge gefeiert. All das verliert an Wert, wenn wir zur Einsicht gelangen, dass das nicht nötig ist. Wir können junge Menschen fördern, ohne sie miteinander zu vergleichen, und ihnen dabei mehr Selbststeuerung und Selbstverantwortung zutrauen. Wir können Lernprozesse ganzheitlich sehen, ohne sie durch Prüfungen künstlich zu zerteilen. Lehrpersonen können mit Schüler:innen Vertrauensverhältnisse aufbauen, in denen sie gemeinsam über Lernen und Leisten nachdenken, um ein *ownership of learning* zu erreichen. Bei der Entfaltung dieser Ideen können Lehrpersonen bereits mit ersten kleinen Maßnahmen wertvolle Impulse in die Richtung einer sinnvollen Lernkultur setzen.

Dieses Buch soll einen Beitrag dazu leisten. Es ist aus der Arbeit des *Instituts für zeitgemäße Prüfungskultur* entstanden, die im letzten Beitrag vorgestellt wird. Entscheidend ist, dass es um einen Kulturwandel geht, der viel mehr umfasst als Digitalisierung. Der Vorstellung, Prüfungen seien dann gut, wenn sie digital abgenommen werden, muss deshalb an dieser Stelle vehement widersprochen werden. Prüfungen sind dann gut, wenn sie Schüler:innen in ihrem Lernen und ihrer Entwicklung stärken. Wenn sie spüren können, wie leistungsfähig sie sind, wie erfolgreich sie sein können, wenn sie mit anderen zusammenarbeiten. Wenn sie Rückmeldungen erhalten, die ihnen Perspektiven eröffnen. Das ist die Vision, welche die Beiträge in diesem Band verbindet. Einige davon sind länger und theoretischer, sie entwickeln Argumente und Leitvorstellungen bei der Entwicklung einer Prüfungskultur. Andere berichten aus der Praxis und stellen erprobte Formate vor, die einer zeitgemäßen Prüfungskultur entsprechen. Nicht alle Schulformen sind hier so vertreten, wie das wünschenswert wäre.

Der Band ist in drei Teile geteilt:

1. Eine Kritik der bestehenden Prüfungskultur
Leitfrage: Warum muss sich etwas ändern?
2. Eine Sammlung von Argumenten für eine zeitgemäße Prüfungskultur
Leitfrage: Was sollten Prüfungen sein?
3. Hilfestellungen beim Erarbeiten von Prüfungen in einer zeitgemäßen Lernkultur
Leitfrage: Wie gelingt der nächste Schritt?

Traditionelle und zeitgemäße Prüfungskultur – eine Übersicht

Der Begriff »zeitgemäß«, der im Titel dieses Bandes steht, ist schwer zu definieren. Er markiert eine Abgrenzung von etwas, was nicht zeitgemäß ist, drückt aber nicht aus, wie diese Abgrenzung genau vorgenommen wird. Dieses Problem ist uns als Herausgeber:innen bewusst, es würde aber auch alle begrifflichen Alternativen wie

»neu«, »offen« etc. betreffen. Um den Begriff zu füllen und zu konkretisieren, haben wir die folgende Übersicht formuliert, die aus einer Impulssammlung entstanden ist. Sie ist auch eine verdichtete Zusammenfassung der Kritik an etablierter Prüfungskultur (linke Spalte) und der Anforderungen an eine sinnvolle Prüfungskultur (rechte Spalte), die im ersten und zweiten Teil dieses Bandes formuliert werden.

	traditionelle Prüfungskultur	zeitgemäße Prüfungskultur
Zugriff auf Leistungen	Leistungen werden von Lehrpersonen abgefragt.	Leistungen werden von Schüler:innen aktiv nachgewiesen.
Leistungsformen	Leistung ist meist Reproduktion oder Reorganisation von deklarativem Wissen.	Leistung umfasst alle Formen von Kompetenzen, auch soziale, kommunikative, metakognitive.
paradigmatisches Format	Lernende beantworten eigenständig geschlossene Aufgaben mit Erwartungshorizont. Lehrende beobachten sie dabei möglichst unbeteiligt.	Lernende erstellen offene Lernprodukte mit Reflexion, auch in Gruppen. Lehrende beobachten sie teilnehmend und ermöglichen <i>peer-assessment</i> und <i>self-assessment</i> .
Unterrichtsorganisation	Lehrpersonen orientieren sich an Plandidaktik im Stundentakt.	Lehrpersonen designen Umgebungen, in denen Projekte stattfinden.
Arbeitsform	Schüler:innen arbeiten isoliert, mit stark beschränktem Zugriff auf Medien (Buchkultur).	Schüler:innen vernetzen sich für die Zusammenarbeit im <i>Open-Media-Modus</i> (Kultur der Digitalität).
Sinn der Prüfung	Die Prüfung muss erfolgen, damit Beurteilung und Bewertung möglich ist.	Lernende erhalten über die Prüfung eine Einschätzung und Anerkennung für ihre Leistungen, Lern- und Entwicklungsprozesse. Die Prüfung ist Gelegenheit, Verantwortung für Entwicklung (individuell oder in Gruppen) zu übernehmen.
zeitliche Struktur	Die Prüfung ist eine Momentaufnahme, die den Lernprozess summativ rhythmisiert/abschließt, Lern- und Prüfphasen sind getrennt.	<i>Constructive Alignment</i> von Lernen und Prüfen Lernen als längerfristiger, formativer Prozess.

	traditionelle Prüfungskultur	zeitgemäße Prüfungskultur
Ziel der Leistungserfassung	Leistung wird erhoben, um Vergleich, Prognose und Normierung zu ermöglichen, »Assessment of Learning«.	Leistung wird erhoben, um eine Diagnose vorzunehmen und Verbindlichkeit im Lernprozess herzustellen, »Assessment for Learning« und »Assessment as Learning«.
soziale Annahmen	Alle Lernenden lösen zum selben Zeitpunkt dieselben Aufgaben auf eine vergleichbare Art und Weise.	Alle Lernenden haben individuelle Kompetenzprofile und bearbeiten unterschiedliche Problemstellungen.
Rolle der Lehrperson	Erstellt Bewertungshorizonte, korrigiert und beurteilt.	Begleitet den Leistungsnachweis formativ und gibt Feedback.
Vorstellung von Lernen	Lernen erfolgt mechanistisch, ist eine Konditionierung bzw. ein Training.	Lernen erfolgt konstruktivistisch, in der Auseinandersetzung mit authentischen Problemen und in der Vernetzung mit anderen Lernenden.
Orientierungspunkt für Prüfungen	Prüfungen ahmen das naturwissenschaftliche Experiment nach, sie sollen im Idealfall messen.	Prüfungen sind einem Gespräch nachempfunden, sie funktionieren wie ein zwischenmenschlicher Dialog.
Verantwortung für Lernen	Lehrende sind verantwortlich dafür, dass Schüler:innen so lernen, dass sie bei Prüfungen gute Ergebnisse erzielen.	Lernende und Lehrende teilen die Verantwortung fürs Lernen und für die Leistungserhebung (<i>ownership of assessment</i>).

Ein herzlicher Dank gebührt den Menschen, die das Buch möglich gemacht haben: Das sind alle Mitglieder des Instituts, die mitgeschrieben, mitgedacht und mitdiskutiert haben und ohne die es den Band nie gegeben hätte. Danke also an Anna Reuter, Axel Krommer, Björn Nölte, David Tepasse, Dejan Mihajlović, Hendrik Haverkamp, Lars Mecklenburg, Lars Zumbansen, Oliver Schmitz, Patricia Drewes, Uta Hauck-Thum und Verena Knoblauch. Enorm wichtig sind aber auch all die Lehrpersonen, die den Mut haben, eine zeitgemäße Prüfungskultur zu leben, die mit ihren Schüler:innen zusammen neue Formate ausprobieren und Erfahrungen teilen. Von ihnen kommt viel Inspiration, viel Innovation – es sind mehr Personen, als wir hier erwähnen können. Und zum Schluss bedanken wir uns bei den Menschen, die uns so unterstützen, dass wir die Energie für dieses Projekt aufbringen konnten.

Christian Albrecht

Aufgaben und Herausforderungen einer zeitgemäßen Prüfungskultur am Beispiel von 4K und KI

Dass das Bildungssystem die traditionelle Prüfungskultur zunehmend infrage stellt, ist eine zaghafte, aber kaum noch zu übersehende Entwicklung. In Bayern experimentieren weiterführende Schulen im Schulversuch »Prüfungskultur innovativ« mit alternativen Aufgabenformaten, in Niedersachsen sind Berichtszeugnisse als Alternative zu den Ziffernzeugnissen geplant, in Baden-Württemberg startete im Schuljahr 2022/2023 an 35 Grundschulen ein Schulversuch zu lernförderlichen Leistungsrückmeldungen, Berlin denkt (zögerlich) darüber nach, Leistungsbewertung KI-integrierend weiterzuentwickeln usw. Was eine zeitgemäße Prüfungskultur dabei im Kern auszeichnet, wird diskutiert. Geprägt ist diese Diskussion unter anderem von Überlegungen, welche kompetenzorientierten Prüfungsformate sich eignen (Gerick/Sommer/Zimmermann 2018), welche Freiheitsgrade dabei von Bedeutung sind (Albrecht 2021; 2022), wie Leistungen beurteilt (Beutel et al. 2022) und welche Verfahren und Qualitätsmaßstäbe hierfür angesetzt werden sollen (Führer et al. 2023), welche Funktionen Prüfungen erfüllen sollen und welche Rolle Noten dabei spielen (Nölte/Wampfler 2021; Winter 2018).

In diesen Zusammenhängen ist oft die Rede von sogenannten 21st Century oder Future Skills, also Kompetenzen bzw. Dimensionen menschlichen Handelns, bei denen davon ausgegangen wird, dass sie im Alltag und in der Berufswelt des 21. Jahrhunderts von Bedeutung sind und längerfristig sein werden. Auch wenn 21st Century Skills die »Fähigkeiten« prominent im Namen tragen, beschränken sie sich nicht auf diese, sondern umfassen ebenso domänenspezifische und überfachliche Wissens- und Verständnisanforderungen sowie Einstellungen, Werte und ethische Verhaltensweisen (Binkley et al. 2012). Mit diesem weiten Begriffsverständnis ist der in diesem Kontext international etablierte Terminus »Skill« als Synekdoche oder »Umbrella Term« (Kotsiou et al. 2022, S. 179) zu verstehen und entspricht eher dem ganzheitlichen Kompetenzbegriff, wie er von der OECD (2019) vertreten wird.

Innerhalb der Vielzahl an Rahmenmodellen, die mittlerweile zu den 21st Century Skills existieren (für einen Überblick siehe Kotsiou et al. 2022; van Laar et al. 2017; Voogt/Roblin 2012), wird den vier »Learning and Innovation Skills« (Tril-

ling/Fadel 2009) Kommunikation, Kollaboration, Kritisches Denken und Kreativität eine Schlüsselrolle zugeschrieben:

»Beyond meeting the new demands of 21st century work, these [...] skills have long been at the heart of what it takes to become a self-reliant lifelong learner. The ability to ask and answer important questions, to critically review what others say about a subject, to pose and solve problems, to communicate and work with others in learning, and to create new knowledge and innovations that help build a better world – these have always been at the core of learning and innovation. So we have both timeless universal reasons and urgent practical needs to put these [...] skill sets on the top of the 21st century skills list« (Trilling/Fadel 2009, S. 49 f.).

Auch wenn diese 4K nicht unumstritten sind – unter anderem wegen unscharfer Begrifflichkeiten, unklarer Systematiken und differierender Schwerpunktsetzungen innerhalb der einzelnen Frameworks, wegen fehlender Trennschärfe zwischen den einzelnen Kompetenzbereichen und mangelnder empirischer Fundierung (Bettinger 2021; Wisniewski/Daumiller 2023; Ehlers 2022, S. 10 f.; Kalz 2023) –, erhielt die Diskussion im deutschsprachigen Raum im Dezember 2021 konkrete bildungspolitische Relevanz, als die Kultusministerkonferenz eine Ergänzung zum Strategiepapier »Bildung in der digitalen Welt« veröffentlichte. In dieser fordert sie explizit die Berücksichtigung der 4K bei der Entwicklung zukünftiger Prüfungsformate (KMK 2021), was die traditionelle Prüfungskultur in ihrem Fundament – der individuellen Leistungsmessung und -bewertung – zu erschüttern vermag.

Zusätzlich disruptiv wirkt der rasante Fortschritt der menschenzentrierten künstlichen Intelligenz, die nun auch (vermeintliche?) anthropologische Spezifika in ihrer Nicht-Automatisierbarkeit infragestellt (siehe auch den Beitrag von Wampfler/Haverkamp sowie den Beitrag von Haverkamp in diesem Band).

Diese jüngsten Entwicklungen stellen die Prüfungskultur des 21. Jahrhunderts vor große Herausforderungen und Aufgaben, denn ein System, das den rasanten technologischen und gesellschaftlichen Entwicklungen Rechnung tragen will, das an der Förderung sinnstiftenden Lernens mehr interessiert ist als an Bewertungen durch Ziffern und für das die 4K konstitutiv sein sollen, muss hierauf Antworten finden. Die folgenden Ausführungen sollen als Ansatzpunkte dienen, um solche Antworten formulieren zu können.

Kreativität

Terminologische Annäherungen

Kreativität ist ein vielschichtiges Phänomen, dem man sich mittels verschiedener Rahmenmodelle der psychologisch ausgerichteten Kreativitätsforschung, insbesondere dem 4P-Modell von Rhodes (1961), dem 4C-Modell von Kaufman und Beghetto (2009) sowie dem 7C-Modell von Lubart (2017), nähern kann. In einem ersten Schritt lassen sich anhand der 4Ps vier grundlegende Leitfragen ableiten: Welcher Typ Mensch ist kreativ (*Persons*)? Was wird als kreativ angesehen (*Product*)? Wie werden wir kreativ (*Process*)? Wie beeinflusst das Umfeld die Entwicklung von Kreativität (*Press*)? Spezifizieren lassen sich diese Leitfragen, wenn man über die verschiedenen Ausprägungen von Kreativität und ihre Entstehungsbedingungen nachdenkt: Kaufman und Beghetto (2009) schlagen hierfür Differenzierungen in vier hierarchischen Abstufungen vor: Auf der höchsten Stufe des Kreativen steht das kreative Genie (*Big-C*), dessen Wirken von herausragender und überdauernder gesellschaftlicher Bedeutung ist. Eine Stufe darunter steht das *Pro-C*, ein professionelles Kreativitätslevel, auf dem etwa die kreativen Leistungen erfolgreicher Schriftsteller:innen, innovativer Wissenschaftler:innen und Sternköch:innen verortet werden können, ohne dass diesen Geniestatus zuerkannt würde. Für schulische Lern- und Bildungsprozesse relevant sind vor allem das sogenannte *mini-c* und das *little-c*. Das *mini-c* definieren die Autoren als

»the novel and personally meaningful interpretation of experiences, actions, and events (Beghetto & Kaufman, 2007). [...]. Central to the definition of *mini-c* creativity is the dynamic, interpretive process of constructing personal knowledge and understanding within a particular sociocultural context« (Kaufman/Beghetto 2009, S. 3).

Für den Bildungskontext und die Diskussion um 21st Century Skills ist vor allem das *mini-c* zentral, da es den Blick auch auf die unauffälligeren, aber subjektiv und persönlich bedeutsamen Einsichten, Entdeckungen und Interpretationen der Schüler:innen in der Auseinandersetzung mit neuen und unbekanntem Themen lenkt, die leicht als unbedeutend abgetan oder übersehen werden können (Kaufman/Beghetto 2009, S. 4). Auf der Stufe des *little-c* äußert sich Kreativität im Vermögen, kreative Lösungen für alltägliche Probleme und Herausforderungen in verschiedenen Lebensbereichen zu finden, die ihrerseits Auswirkungen haben auf das Individuum, aber auch auf dessen unmittelbare Umwelt.

Um individuelle und subjektiv bedeutsame *mini-c*-Ideen in *little-c*-Beiträge zu transformieren, die auch von anderen als kreativ angesehen werden, erachten die

Autoren das Feedback als wichtigstes Unterstützungsverfahren. Um vom little-c auf das Pro-C-Level zu gelangen, das heißt Alltagskreativität auf eine professionelle Ebene zu heben, sei vor allem konsequentes Üben notwendig. Der Big-C-Status schließlich sei nur wenigen Menschen vorbehalten und spiele für schulische Kontexte in der Regel keine Rolle, da er in der Regel erst nach langer Zeit (mitunter erst nach dem Tod der Schöpfer:innen selbst) erreicht werde (Beghetto/Kaufman 2014, S. 55).

Mit Bezug auf die 7Cs von Lubart (2017) lassen sich schließlich kollaborative Anteile am kreativen Prozess (Collaborations) sowie Fragen der Ausbildung und Entwicklung von Kreativität in Bildungskontexten (Curricula) ergänzen.

Entgegen vieler traditioneller Annahmen kann Kreativität auf unterschiedliche Weise gefördert werden, sowohl durch direktes, aktives Lehren von Kreativitätskonzepten und -techniken als auch indirekt durch die Schaffung kreativitätsfördernder Lernumgebungen (Beghetto 2019; Beghetto/Kaufman 2014; Chiu 2015; Thornhill-Miller/Dupont 2016; Thornhill-Miller et al. 2023).

Kreativität als 21st Century Skill

Vor dem Hintergrund der 21st Century Skills ist Kreativität vor allem als kreatives Denken zu verstehen, das dabei helfen soll, innovative Lösungen für die komplexen Probleme und Herausforderungen der Gegenwart und Zukunft zu finden. Kreativität gilt als unverzichtbares Element für die Flexibilität und Anpassung des Einzelnen in schwierigen Situationen (Sternberg 1986), als »the most important economic resource of the 21st century« (Florida 2002, zit. n. Kaufman/Beghetto 2009, S. 1) und als »important skill that can help students navigate uncertain futures« (Beghetto/Kaufman 2014, S. 53).

Binkley et al. (2012, S. 37 ff.) führen als notwendiges Wissen z. B. Kenntnisse über Kreativitätstechniken sowie ein historisches Verständnis für kreative Leistungen und deren Grenzen an. Auf der Ebene der Skills verorten sie die Fähigkeit, neue Ideen sowohl allein als auch kollaborativ zu entwickeln, auszuarbeiten, zu verfeinern und zu reflektieren. Hierzu gehört auch die Fähigkeit, Ideen zielgerichtet zu kommunizieren und dabei sensibel für historische und kulturelle Grenzen bei der Durchsetzung kreativer Ideen zu sein. Beharrlichkeit, Durchhaltevermögen, die Offenheit für verschiedene Perspektiven, eine positive Fehlerkultur sowie die Fähigkeit zu konstruktivem Feedback sind wichtige Einstellungen und Haltungen für die Entwicklung von kreativem Denken.

Kreativität, Künstliche Intelligenz und Prüfungskultur

Kreative Prüfungsformate müssen Aufgaben beinhalten, die individuelle, innovative und kreative Lösungen fordern und fördern. Geschlossene Aufgaben, die nur

einen oder wenige Bearbeitungswege zulassen, so wie es in der traditionellen Prüfungskultur überwiegend der Fall ist, eignen sich dagegen nicht. Es gilt Prüfungsformate zu etablieren, in denen die Schüler:innen verschiedene Kreativitätstechniken anwenden können, mit deren Hilfe sie zur individuellen und kollaborativen Auseinandersetzung mit überfachlichen und fachlichen Problemen motiviert werden und in denen sie Lösungen für diese Probleme auf inhaltlich, formal und medial vielfältige Weise umsetzen, präsentieren und reflektieren können. Da sich eine kreativitätsfördernde Lernumgebung durch eine positive Fehlerkultur sowie eine angstfreie Lernatmosphäre auszeichnet (siehe den Beitrag von Schmitz/Haverkamp in diesem Band), eignen sich vor allem projektorientierte, formative Assessments als Rahmen für die Entwicklung zeitgemäßer Prüfungsformate, die auch der Reflexion über Misserfolge und den daraus gewonnenen Erkenntnissen Raum geben.

Auch künstliche Intelligenz kann in kreative Prüfungsformate sinnvoll integriert werden, etwa wenn die Aufgabe explizit auf das explorativ-kreative Spiel mit künstlicher Intelligenz abzielt. Die kreative Leistung könnte dann unter anderem im Prompt Design bzw. Prompt Engineering bestehen, indem Schüler:innen Anweisungen an eine KI sprachlich so gestalten und optimieren, dass sie bestmögliche Ergebnisse für das zu lösende Problem liefert.

»Das verlangt von Lernenden aber auch, sich bereits vor der Nutzung ganz grundlegend mit ihren Interessen und Fragestellungen auseinandergesetzt zu haben und diese formulieren zu können. Prompting ist sicher Bestandteil einer neuen Schreibkompetenz, wenngleich unklar bleibt, ob mangelnde Fähigkeiten nicht auch von den Systemen kompensiert werden können, indem z. B. Prompts vorgeschlagen werden« (Schindler 2023, S. 15).

Der Einsatz von KI in kreativen Prüfungsformaten sorgt schließlich dafür, dass sich die Anforderungsbereiche – oder die Grenzen »zwischen dem, was als kreativ, und dem, was als mechanisch gilt« (Stalder 2017, S. 176) – verschieben: Produktive, unter Umständen auch künstlerisch-handwerkliche Fähigkeiten könnten an Bedeutung verlieren, während ästhetische Reflexionsfähigkeit und ästhetische Urteilskompetenz – und vielleicht sogar ästhetische Interpretationsfähigkeit in Bezug auf künstliche Kreativität – an Relevanz gewinnen. Die Auseinandersetzung mit den Ergebnissen algorithmischer Kreativität kann zudem den Blick und den Sinn schärfen für die emotionalen, motivationalen, kulturellen und persönlichen, das heißt genuin menschlichen Implikationen kreativen Denkens sowie für Fragen zu Originalität, Autorschaft und Authentizität.

Die Frage nach der Beurteilung kreativer Leistungen offenbart jedoch ein Objektivitäts-Validitäts-Dilemma: Denn kreative Leistungen können nur dann objektiv bewertet werden, wenn die Rahmenbedingungen für Schüler:innen konstant und vergleichbar gehalten werden. Dies geht allerdings auf Kosten der Validität, da

das, »was einigermaßen objektiv erhoben und beurteilt werden kann, [...] die Schülerleistung nur noch sehr ausschnitthaft und unvollständig [repräsentiert]« (Sacher 2009, S. 48). Die Beurteilung kreativer Leistungen macht also eine Anpassung der Maßstäbe erforderlich: Anstelle von Objektivität und Reliabilität könnten dann die Kriterien Nützlichkeit, Neuheit und Nachhaltigkeit treten.

Kritisches Denken

Terminologische Annäherungen

Kritisches Denken berührt viele verschiedene Kompetenzfacetten und wird in unterschiedlichen Diskursen unterschiedlich konturiert (Thornhill-Miller et al. 2023). Eine Basisdefinition liefern Pasquinelli und Kolleg:innen:

»critical thinking is the capacity of assessing the epistemic quality of available information and – as a consequence of this assessment – of calibrating one’s confidence in order to act upon such information« (Pasquinelli et al. 2021, S. 169).

Weitgehend Konsens besteht hinsichtlich der sechs Core Critical Thinking Skills Interpretation, Analysis, Inference, Evaluation, Explanation und Self-Regulation sowie der entsprechenden Werte, Einstellungen und Verhaltensweisen, die Facione (2023) im Rahmen einer Delphi-Studie identifiziert hat. Er resümiert:

»The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit« (Facione 2023, S. 26).

Kritisches Denken als 21st Century Skill

Die Vorstellung von kritischem Denken im Rahmen der 21st Century Skills fokussiert das sinnvolle Argumentieren, systemisches Denken, Problemlösungskompetenzen sowie die entsprechenden Einstellungen, Werte und Haltungen. Binkley et al. (2012) subsumieren hierunter die Fähigkeiten, sowohl induktiv als auch deduktiv situationsangemessen argumentieren und dabei unterschiedliche Argumente beurteilen und hinterfragen zu können, diese zusammenzuführen, zu interpretie-

ren und begründet Schlussfolgerungen zu ziehen. Gleichzeitig sollen Schüler:innen alternative Standpunkte in Betracht ziehen und eigene Denkprozesse erläutern und darstellen können. Auf der Ebene der Attitudes/Values/Ethics ergänzen Binkley und Kolleg:innen zudem die Offenheit für ungewohnte und innovative Problemlösungen, das Vertrauen in die Vernunft, Wissbegierde, Flexibilität, Ehrlichkeit und Fairness sowie die Bereitschaft, eigene Ansichten zu überdenken.

Kritisches Denken, Künstliche Intelligenz und Prüfungskultur

Für eine zeitgemäße Prüfungskultur bedeutet diese Schlüsselkompetenz eine Abkehr von Reproduktionsaufgaben, deren Antworten auch im Internet recherchiert und per Copy/Paste eingesetzt werden könnten. Stattdessen gewinnen Aufgaben an Bedeutung, in denen die Schüler:innen ihre Fähigkeit zum konzeptionellen, kreativen, synthetisierenden und eben kritischen Denken unter Beweis stellen können. Hierfür eignen sich problemlösende und Transferaufgaben, die gerade die kritische Recherche nach und Auseinandersetzung mit Informationen sowie Such- und Bewertungsstrategien erfordern. Zum Lösungsweg gehört ebenfalls die Diskussion, online wie offline, mit anderen Noviz:innen und Expert:innen. Dass der Prüfungsgegenstand dabei nicht länger monomedial auf gedruckte Medien beschränkt sein sollte, sondern auch digitale schriftliche, auditive, visuelle, audiovisuelle und symmediale Texte Gegenstand kritischer Reflexion sein müssen, liegt angesichts der zunehmenden Präsenz von Desinformation und Propaganda, von inhaltlichen Verzerrungen und Falschmeldungen insbesondere in digitalen Kontexten nahe. Schüler:innen müssen die Gelegenheit bekommen, in Prüfungen mit verdeckten und offenen Intentionen von Texten und weiteren Medien umzugehen und deren Wahrheitsgehalt kritisch zu erfassen. Dies schließt »das Wissen um Mechanismen selektiver und adressatenspezifischer emotionaler Manipulation mit ein« (Frederking/Krommer 2019, S. 9). Auch mündliche Prüfungsformate wie die Debatte gewinnen in diesem Sinne an Bedeutung.

Nicht zuletzt die rasante Entwicklung künstlicher Intelligenz macht das kritische Denken zu einer zentralen Bildungsaufgabe im 21. Jahrhundert, da nun ohne größeren Aufwand viele Nachrichten im großen Stil fingiert, KI-generierte Video- und Audio-Dateien produziert, Social-Media-Kommunikation durch Bots polarisiert und Medien in Text, Bild und Ton manipuliert werden können. Andererseits erfordern die (momentan noch) fehleranfälligen, mitunter auch »halluzinierten« automatisierten Ergebnisse künstlicher Intelligenz selbst einen kritischen Umgang. Für Prüfungssituationen bedeutet dies, künstliche Intelligenz nicht auszuschließen, sondern sie zu integrieren und den Umgang mit ihr zu reflektieren und zu erläutern. Als zentraler Maßstab dabei gilt der Grundsatz wissenschaftlicher Redlichkeit: Quellen und Prompts werden vollständig genannt, indirekte und direkte Zitate sowie Paraphrasen kenntlich gemacht, fremde Gedan-



Probleme der traditionellen Prüfungskultur

Philippe Wampfler

Meritokratie und individuelle Leistung als Märchen für Erwachsene

In ihrem Buch *Erzählende Affen* (2020) beschreiben Samira El Ouassil und Friedemann Karig, welche Bedeutung Narrative für das menschliche Leben haben. Als »Märchen für Erwachsene« bezeichnen sie dabei Erzählungen, die »Herrschaft und Ungerechtigkeit verklären können, so glaubhaft, dass wir nicht mehr bemerken, dass die Grundlage unserer Gesellschaften elaborierte Fiktionen über die Ordnung der Dinge sind« (El Ouassil/Karig 2020, S. 234).

Schule und Bildung sind überlagert von einer Reihe dieser Märchen für Erwachsene. Lehrpersonen und Verantwortliche in der Bildungspolitik erzählen sie, um Bildungsungerechtigkeit auszublenken und problematische Praktiken zu stützen und zu rechtfertigen. Eines dieser Märchen betrifft die Meritokratie. El Ouassil und Karig schreiben dazu:

»Erfolg und Misserfolg als Ergebnis von Fleiß und Talent. So akzeptierten wir zunehmend, dass manche Menschen aufsteigen und andere darben, weil wir dies mit Anstrengung und Leistung begründen konnten. Die neue Herrschaft ging nicht mehr von einem Menschen qua Geburt aus, sondern von einem Menschen, der sich seine Position »verdient«. Auf diesem Narrativ der Meritokratie – Erfolg durch Arbeit – basiert die gesamte ökonomische und soziale Selbsterzählung der westlichen Leistungsgesellschaften. Sie ist heute vielleicht das wirkmächtigste Märchen für Erwachsene überhaupt. [...] Die meritokratische Auffassung erkennt, dass die Zugänge zu Bildung, Berufen und Beziehungen nicht allen gleichermaßen zur Verfügung stehen.

Bereits die Schulbildung sorgt für soziale Ungleichheiten, die sich nicht ohne Weiteres aus Begabung, Leistung oder Talent ableiten lässt. In nahezu allen westlichen Gesellschaften ist der sicherste Indikator für Erfolg in Bildung und Beruf immer noch der Status der Eltern. Erfolg scheint so gesehen weniger erworben als vererbt zu werden. Besser qualifizierte Menschen nehmen dennoch in der Regel an, dass sie ihre goldene Zukunft ihren Leistungen verdanken. [...] Der narrative Trick hierbei ist, die vom gesellschaftlichen Tisch Ausgeschlossenen davon zu überzeugen, dass sie ganz allein für ihren Ausschluss verantwortlich sind. Mehr noch: dass sie denen, die etwas leisten, auf der Tasche liegen« (El Ouassil und Karig 2020, S. 256).

Die Vorstellung, schulische Leistungen rechtfertigten es, Schüler:innen unterschiedlichen Bildungswegen und Berufsmöglichkeiten zuzuteilen, ist ein wirksames Narrativ, das die Ungerechtigkeit einer solchen Diskriminierung überdeckt: Die Annahme, Menschen, die im Bildungssystem diskriminiert werden (weil sie weniger lang in die Schule dürfen, weil ihre Lehrkräfte schlechter bezahlt sind, weil die Schulräumlichkeiten weniger gut ausgestattet sind etc.), hätten das aufgrund ihrer Leistungen verdient, ist ethisch und pädagogisch nicht haltbar. Kinder haben Zurücksetzungen nie verdient – sie haben alle denselben Anspruch auf Lerngelegenheiten und Förderung. Das meritokratische Denken ist ideologisch deshalb wichtig, weil Bildungseliten das Schulsystem benutzen, um ihre Privilegien weiterzugeben. Alle Kinder und Jugendlichen verdienen grundsätzlich dieselben Möglichkeiten zum Lernen. Da Bildung dafür missbraucht wird, Privilegien zu generieren, werden angebliche »individuelle Leistungen« vorgeschoben, um zu begründen, weshalb Bildung ungerecht sein müsse. Der Wert von Bildung ist nicht an Leistung geknüpft: Alle Menschen profitieren davon.

Der Bildungsforscher Krassimir Stojanov beschreibt in seinem Buch *Bildungsgerechtigkeit* (2011) die negativen Auswirkungen des meritokratischen Schulmärchens folgendermaßen:

»Die Folge dieser planwirtschaftlich ausgerichteten Forderung für die Betroffenen ist, dass sie schon vor ihrem eigentlichen Eintritt in das Arbeitsleben als Menschen mit bestimmten Kompetenzen, Fähigkeitspotenzialen oder Begabungen festgelegt werden. Diese Festlegung von außen unterdrückt in vielen Fällen das Selbstwertgefühl und letztlich die Autonomiefähigkeit des Einzelnen. Es verursacht schließlich ein viel größeres psychisches Leiden, sehr früh als jemand etikettiert zu werden, dessen Leistungsfähigkeiten und »Begabung« nur für eine Berufsposition mit einem niedrigeren Status ausreichen, als eine solche Position später aufgrund von zufälligen Umständen im Arbeitsmarkt, oder aufgrund von eigenverantwortlichen Entscheidungen als Erwachsener einnehmen zu müssen« (Stojanov, 2011, S. 106).

Meritokratie ist kein Ideal, das diejenigen belohnt, die das verdient haben. Sie ist eine Dystopie, die vom System Benachteiligten einredet, dafür selber verantwortlich zu sein. Michael Sandel hat in *The Tyranny of Merit* (2020) gezeigt, dass die Vorstellung, Erfolg und Misserfolg seien auf bewusste Handlungen von Menschen zurückzuführen, ausblendet, wie stark Umstände und Glück tatsächlich wirken. Ursprünglich wurde der Begriff denn auch erfunden, um diese Dystopie zu kennzeichnen. Im Rahmen einer politischen Ideologisierung hat er aber eine positive Prägung erfahren, die heute die Verteilung beschränkter Ressourcen in vielen Gesellschaften prägt.

Meritokratie ist der Rahmen, in dem viele Probleme, die mit einer traditionellen Prüfungskultur verbunden sind, verortet werden können. Sie alle orientieren sich an der Vorstellung, Schule müsse und könne Leistungen von Lernenden messen, bewerten und vergleichen. Dabei blenden sie aus, dass diese Vorgänge zu Daten führen, die im Bildungssystem dazu genutzt werden, um Schüler:innen zu diskriminieren. Die Leistungsfeststellung ist immer mit Selektion und Allokation verbunden (vgl. dazu das entsprechende Kapitel in diesem Band). Wäre das nicht der Fall, wäre sie gar nicht nötig – Leistungen müssten in der Schule weder gemessen noch verglichen werden. Dies mag für viele utopisch klingen, ist es aber dann nicht, wenn Allokation und Selektion aus der Schule ausgelagert würden. Das hätte den Vorteil, gesellschaftlich transparent zu machen, dass diese Vorgänge ablaufen, statt sie in der Schulsozialisierung zu normalisieren.

Formatives Feedback, das Lernenden Hinweise darauf gibt, wie ihre Lernprozesse ablaufen und wie sie optimiert werden können, wäre in einer Schule, die ungerechten Bildungsprozessen nicht zudient, die einzig nötige Form von Rückmeldung. Summative Bewertungen wären nicht nötig.

Die Trennung von Lernen und Leistung

Im Rahmen der Kompetenzorientierung haben Didaktiker:innen wie Josef Leisen, Juliane Köster oder Ulf Abraham eine Differenzierung von zwei Aufgabentypen vorgenommen: Sogenannte »Lernaufgaben« dienen dazu, Wissen zu erwerben und Kompetenzen aufzubauen, während sogenannte »Leistungsaufgaben« dazu konzipiert sind, Wissenserwerb oder Kompetenzen zu überprüfen (für eine Übersicht vgl. Luthiger 2012, S. 2 f.). Diese Unterscheidung hatte gravierende Konsequenzen für die Ausbildung von Lehrkräften: Diese lernen, Unterricht in zwei Modi aufzuteilen. Im Kompetenz-Modus liegt der Fokus auf der Entwicklung der Lernenden, auf ihrem Lernen. Dafür werden Lernaufgaben eingesetzt. Im Performanz-Modus hingegen steht der Status im Vordergrund, Lernende kämpfen um einen Rang in der Klasse, ihr Ziel ist nicht eine Entwicklung, sondern symbolisches Kapital in der Form von Noten. Diese beiden Modi lassen sich auf eine bildungssoziologische Doppelaufgabe von Lehrpersonen zurückführen (Luthiger 2012, S. 6 f.). Einerseits müssen sie Lernende fördern, andererseits müssen sie prüfen, um Allokation und Selektion betreiben zu können. Das ist ein fundamentaler Widerspruch im Rollenbild. Statt diesen Widerspruch aufzulösen, indem die meritokratische Aufgabe der Belohnung und Bestrafung je nach Leistung zurückgewiesen wird, schafft die begriffliche und funktionale Trennung von Lern- und Leistungsphasen im Unterricht eine Möglichkeit, die unterschiedlichen psychologischen Dynamiken von Lern- und Prüfungssituationen zu legitimieren.

Werden beispielsweise Prüfungsaufgaben kritisiert, weil sie zu wenig offen oder nicht authentisch sind, dann lässt sich aus der Perspektive des Performanz-Modus argumentieren, es handle sich eben um »Leistungsaufgaben«, die gar nicht die Funktion hätten, offene Lösungswege oder authentische Problembearbeitungen zuzulassen. Wie bei der Meritokratie handelt es sich hier auch um ein Märchen für Erwachsene: Es blendet aus, dass die problematischen Prüfungsbedingungen Aufgabenformate erzwingen, die wesentliche Arbeitsformen aus dem Lernalltag und aus dem Unterricht verunmöglichen. Viele Lehrpersonen glauben an diese Erzählung, weil sie nicht akzeptieren können, wie problematisch ihre Beurteilungsfunktion für Schüler:innen ist und wie stark sie mit Bildungsgerechtigkeit verbunden ist.

Viele der Beiträge und Argumentationslinien in diesem Band verschieben den Fokus hin zu Lernaufgaben. Sie versuchen, Prüfungssituationen an den Kompetenz-Modus anzubinden, in dem Schüler:innen sich entwickeln können. Der Blick auf die Diskussion um die Unterscheidung von Lern- und Leistungsaufgaben kann hier deshalb hilfreich sein, weil sie zeigt, dass unter Umständen falsche Erwartungen erzeugt werden: Werden Noten generiert, dann arbeiten Schüler:innen anders, als wenn sie sich entwickeln dürfen, ohne dafür bewertet zu werden (Luthiger 2023, S. 7). Diese psychologische Einsicht kann nicht unterlaufen werden, indem die Unterscheidung von Lern- und Leistungsaufgaben unsichtbar gemacht wird oder Lernaufgaben dazu benutzt werden, Leistungen zu erheben.

Solange Menschen an individuelle Leistung und Meritokratie glauben, ist es schwierig, eine Schule zu gestalten, in der diese Gedanken die Prüfungskultur nicht beeinflussen. Um zu verstehen, wo neben einer meritokratischen Gesellschaft das Problem bei der Leistungsorientierung liegt, braucht es noch einen weiteren Gedanken, der den Mythos zerstört, Menschen könnten individuelle Leistungen erbringen.

Die Erfindung der Leistung

In ihrem Buch *Die Erfindung der Leistung* von 2018 hat Nina Verheyen ausführlich dargelegt, weshalb es keine isoliert erbrachten Lernerfolge geben kann: Leistung gründet erstens auf sozialen Parametern, die nicht aus dem Individuum hervorgehen (Verheyen 2018, S. 199 f.). Einige Schüler:innen können von Voraussetzungen profitieren, die es in einer Prüfungssituation so erscheinen lassen, als würden sie eine bessere Leistung als andere erbringen – was aber nur der Fall ist, weil alle sozialen Bedingungen dieser Leistung ignoriert werden. Zweitens entsteht Leistung in einem sozialen Kontext. Schüler:innen lernen von ihren Eltern, von Peers, von Lehrpersonen. Was sie können, ist direkt daran geknüpft, mit wem sie in einer Beziehung stehen. Die Vorstellung, eine Leistung könne über eine Prüfung einem Individuum zugeschrieben werden, ist deshalb mindestens problematisch, wenn nicht absurd.