

Bei dieser Arbeit handelt es sich um eine Wissenschaftliche Hausarbeit, die an der Universität Kassel angefertigt wurde. Die hier veröffentlichte Version kann von der als Prüfungsleistung eingereichten Version geringfügig abweichen. Weitere Wissenschaftliche Hausarbeiten finden Sie hier: <https://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2011040837235>

Diese Arbeit wurde mit organisatorischer Unterstützung des Zentrums für Lehrerbildung der Universität Kassel veröffentlicht. Informationen zum ZLB finden Sie unter folgendem Link:

www.uni-kassel.de/zlb

Wissenschaftliche Hausarbeit

im Rahmen der Ersten Staatsprüfung für das Lehramt an Gymnasien im Fach
Erziehungswissenschaften, eingereicht der Hessischen Lehrkräfteakademie

- Prüfungsstelle Kassel -

Thema:

**Die Bewertung mündlicher Mitarbeit – eine empirische Studie zu
Anspruch und Wirklichkeit der schulischen
Bewertungspraxis**

Gutachter: Prof. Dr. Martin Hänze

Verfasserin: Alina Mangold

Studiengang: Lehramt an Gymnasien (L3)

Datum: 07.11.2016

Inhalt

1. Einleitung und Problemstellung.....	1
2. Rechtliche Vorgaben.....	4
3. Theoretischer Hintergrund	10
3.1 Leistungsbegriffe im schulischen Kontext.....	10
3.1.1 Leistung	10
3.1.2 Schulleistung	12
3.1.3 Leistungsprinzip	14
3.1.4 Zeugnisse und Noten.....	16
3.1.5 Mündliche Leistungen	18
3.1.6 Mündliche Mitarbeit.....	21
3.2 Leistungen feststellen, beurteilen und bewerten.....	24
3.2.1 Leistungsfeststellung	25
3.2.1.1 Objektive Verfahren.....	25
3.2.1.2 Gütekriterien.....	26
3.2.1.3 Subjektive Verfahren	27
3.2.2 Leistungsbeurteilung und -bewertung	29
3.2.2.1 Bezugsnormen	30
3.2.2.2 Herausforderungen beim Beurteilen und Bewerten	31
3.3 Mündliche Mitarbeit feststellen, beurteilen und bewerten.....	33
3.3.1 Kriterien	33
3.3.2 Chancen und Grenzen.....	37
3.3.2.1 Vorteile gegenüber der Bewertung schriftlicher Leistungen	37
3.3.2.2 Herausforderungen	37
3.3.3 Forschungsstand	40
3.4 Zusammenfassung und Ausblick.....	45

4. Studie „Die Bewertung mündlicher Mitarbeit im Unterricht“	47
4.1 Forschungsfragen	47
4.2 Methode	48
4.2.1 Stichprobe	48
4.2.2 Design und Durchführung	50
4.3 Ergebnisse	52
4.3.1 Kriterien	52
4.3.1.1 Benannte Kriterien.....	52
4.3.1.2 Stellenwert der Kriterien	54
4.3.1.3 Stellenwert der Kriterien nach Aufgabenfeldern.....	56
4.3.1.4 Stellenwert der Kriterien nach Berufserfahrung	63
4.3.1.5 Schulische Vorgaben.....	67
4.3.1.6 Absprachen im Kollegium.....	68
4.3.1.7 Bezugsnormorientierung	68
4.3.2 Dokumentation.....	71
4.3.3 Reflexion.....	73
4.4 Diskussion.....	75
4.4.1 Zusammenfassung der Ergebnisse.....	75
4.4.2 Limitation der Ergebnisse und des Forschungsdesigns	78
5. Fazit.....	78
6. Literaturverzeichnis.....	83
7. Abbildungsverzeichnis	89
8. Tabellenverzeichnis	90
9. Schlusserklärung	92
10. Anhang	93

1. Einleitung und Problemstellung

Die mündliche Mitarbeit ist eine geforderte Leistung, mit der alle Lernenden an allgemeinbildenden Schulen im Verlauf der Schulzeit konfrontiert werden. Als bedeutender Teil der „mündlichen Note“ kann die mündliche Mitarbeit den Schulerfolg massiv beeinflussen. Je nach Unterrichtsfach bestimmt die mündliche Note in Hessen die Hälfte bis zwei Drittel der Gesamtnote (vgl. Hessisches Kultusministerium 2006, S. 21). Sie bietet dadurch für Lernende die Gelegenheit, „verpatzte“ schriftliche Leistungen auszugleichen. „Du solltest Dich mehr am Unterricht beteiligen“ ist ein üblicher Appell seitens der Lehrkräfte. Aber was macht eine gute mündliche Beteiligung im Unterricht aus? Wie und anhand welcher Kriterien wird die mündliche Mitarbeit bewertet? Darauf scheint es keine einheitlichen Antworten zu geben. Deshalb müssen sich Schülerinnen und Schüler (im Folgenden durch SuS abgekürzt) auf jede Lehrkraft neu einstellen.



Abbildung 1: Karikatur aus Souvignier & Förster (2014, S. 7)

Die mündliche Note kann vielfältige (mündliche und schriftliche) Leistungen beinhalten. Daher wird sich in dieser Arbeit auf die mündliche Mitarbeit, als elementarer Bestandteil der mündlichen Note, beschränkt. Hierfür werden zuerst

die rechtlichen und theoretischen Hintergründe der mündlichen Mitarbeit beleuchtet. Dabei wird der Fokus auf dem Anspruch an schulische Leistungsfeststellung und -bewertung sowie den Chancen und Grenzen der Bewertung mündlicher Mitarbeit liegen. Auf dieser Grundlage wird mithilfe einer eigenen empirischen Studie untersucht, wie die Bewertung in der Praxis aus der Perspektive der Lehrkräfte umgesetzt wird.

Zunächst werden in *Kapitel 2* die rechtlichen Vorgaben zur Leistungsbewertung dargestellt und auf diese Weise grundlegende Anforderungen an schulische Leistungsbewertung definiert. Dabei wird sich weitestgehend auf hessische Vorgaben beschränkt. Zudem werden die erheblichen Freiheiten der Lehrkräfte bei der Bewertung mündlicher Mitarbeit aufgezeigt.

Anschließend werden in *Kapitel 3* die theoretischen Grundlagen herausgearbeitet. Dafür soll zunächst einmal definiert werden, was den Leistungsbegriff bei verschiedenen Zugängen sowie im schulischen Kontext auszeichnet. Weiterhin wird dargestellt, wie sich durch die Entstehung des Leistungsprinzips in der Gesellschaft ein Bedarf nach der Bewertung von Schulleistungen mithilfe von Zeugnissen und Noten entwickelt hat. Die in der Schule erbrachten Leistungen beeinflussen demnach die zukünftigen Lebenschancen der SuS. Die Kritik an Leistungsbewertung mithilfe von Noten im Allgemeinen wird kein Teil dieser Arbeit sein.

Um zu verstehen, was die mündliche Mitarbeit den SuS an Leistung abverlangt, werden ferner die Unterschiede zwischen Mündlichkeit und Schriftlichkeit und damit zwischen mündlichen und schriftlichen (Schul-) Leistungen dargelegt. Dazu wird grundlegend auf die Verschiedenheiten zwischen mündlicher und schriftlicher Kommunikation eingegangen. Des Weiteren werden die für mündliche Leistungen erforderlichen Fähigkeiten und Fertigkeiten beleuchtet.

Anschließend wird die in der Literatur zugunsten mündlicher Prüfungen eher vernachlässigte mündliche Mitarbeit definiert und von anderen mündlichen Leistungen und dem Begriff der mündlichen Note abgegrenzt.

Bevor auf die Bewertung mündlicher Mitarbeit eingegangen werden kann, wird zunächst der Prozess von der erbrachten Leistung bis zur Note dargestellt. Hierbei wird in der pädagogischen Diagnostik meist zwischen zwei Phasen unterschieden, der Leistungsfeststellung und Leistungsbewertung. Diese Unterscheidung wird in der Gliederung dieser Arbeit aufgegriffen und durch eine weitere Unterscheidung zwischen Beurteilen und Bewerten ergänzt. Da es für den gesamten Prozess keinen einheitlichen Begriff gibt, wird er in dieser Arbeit aus Gründen der besseren Lesbarkeit als „Leistungsbewertung“ bezeichnet.

Hinsichtlich der Leistungsfeststellung werden sogenannte „objektive“ und „subjektive“ Verfahren sowie in diesem Kontext die messtheoretischen Gütekriterien dargestellt. Im Bezug zur Leistungsbeurteilung und -bewertung werden die möglichen Bezugsnormen erläutert. In diesem Abschnitt werden dementsprechend theoretische Ansprüche an schulische Leistungsbewertung herausgearbeitet. Zudem werden mögliche Herausforderungen beim Beurteilen und Bewerten benannt.

Im letzten Abschnitt des theoretischen Teils werden zuerst mögliche Kriterien zur Bewertung mündlicher Mitarbeit erarbeitet. Hierbei werden verschiedene Ansätze und strittige Aspekte diskutiert. Dann werden Vorteile gegenüber der Bewertung schriftlicher Leistungen, aber auch einige Herausforderungen dargestellt. Zum Schluss werden die Ergebnisse der wenigen, bisher durchgeführten Studien zum Thema mündliche Mitarbeit bzw. mündliche Note vorgestellt. Diese behandeln u. a. die Kenntnis über rechtliche Vorgaben, Methoden und Häufigkeit der Dokumentation der Leistungen, die Kriterien sowie die Nennungen von Problembereichen bzw. der Akzeptanz der SuS. Hierbei werden in einer Studie von Krieger (2003) die Angaben von Lehrerinnen und Lehrern (im Folgenden durch LuL abgekürzt) und die der ehemaligen SuS gegenübergestellt.

Im Anschluss werden daraus für den in *Kapitel 4* dargestellten empirischen Teil Forschungsfragen für die eigene populationsbeschreibende Studie abgeleitet, an der LuL aus Kassel und Umgebung teilgenommen haben. Hierbei wird ein Schwerpunkt auf dem unterschiedlichen Stellenwert einzelner Kriterien zur Bewertung mündlicher Mitarbeit bei verschiedenen Schulfächern und Lehrkräften mit unterschiedlich langer Berufserfahrung liegen. Weiterhin wird untersucht, ob die früheren Ergebnisse hinsichtlich der Dokumentation und die Einschätzung über die Zuverlässigkeit der Bewertung bestätigt werden können. Außerdem werden die LuL befragt, ob sie, unabhängig von den rechtlichen Vorgaben, Vorschriften durch die Schule oder Hilfestellungen bzw. Absprachen im Kollegium erhalten. Nach der Vorstellung des Forschungsdesigns und der Stichprobe werden die Ergebnisse der Befragung dargelegt.

Ferner werden diese Ergebnisse im Diskussionsteil zusammengefasst. Dabei werden außerdem die aufgrund des Forschungsdesigns entstandenen Einschränkungen über deren Aussagekraft auf die Grundgesamtheit dargelegt.

Im abschließenden Teil der Arbeit wird vor dem Hintergrund der theoretischen Grundlagen und aus den Ergebnissen der Studie ein Fazit gezogen. Dazu werden die Ansprüche an die schulische Leistungsbewertung der, durch die Studie festgestellten, Wirklichkeit gegenübergestellt. Weiterhin wird ein Ausblick auf

weiteres Forschungspotential gegeben und mögliche Konsequenzen für die Bewertung mündlicher Mitarbeit aufgezeigt.

2. Rechtliche Vorgaben

In diesem Kapitel werden zunächst elementare Grundsätze des deutschen Schulrechts bezüglich der Leistungsbewertung dargestellt. Danach folgen die rechtlichen Grundlagen des Landes Hessen zur Leistungsbewertung, indem zuerst die wesentlichen Passagen aus dem Hessischen Schulgesetz (HSchG) beschrieben und anschließend die dazugehörigen Verordnungen zum Hessischen Schulgesetz und der Oberstufen- und Abiturverordnung (OAVO) dargelegt werden.

Nach Avenarius & Füssel (2010) sind bei der Leistungsbewertung die tatsächlich erbrachten Leistungen der SuS zu berücksichtigen und keine sachfremden Aspekte¹. Zudem muss die Leistung einem Lernenden eindeutig zugeordnet werden können. Bei der Bewertung von Gruppenarbeiten muss beispielsweise die individuelle Leistung des Lernenden erkennbar sein.

Schulische Leistungsbewertung ist ein Prozess, bei dem eine Lehrperson ein persönliches Fachurteil in einem bestimmten Bezugsrahmen trifft. Hierbei wird der Lehrkraft ein gewisser Beurteilungsspielraum zugestanden. Die Bewertung der Lernenden soll zwar die individuelle Lernentwicklung berücksichtigen, sich jedoch in erster Linie an generellen Maßstäben, wie Bildungsstandards und Lehrplänen, orientieren. Grundsätzlich soll bei der Leistungsbewertung eine Gleichbehandlung und Chancengleichheit der Lernenden nach Art. 3 Abs. 1 des Grundgesetzes gegeben sein. Weisen schulische Bewertungen Mängel im Prüfungsverfahren auf, wie z. B. unterschiedliche Bedingungen bzw. Hilfsmittel oder die Befangenheit einer Lehrperson, können diese ggf. vor Gericht angefochten werden (vgl. Avenarius & Füssel 2010, S. 445 ff.).

Für die Beurteilung und Bewertung mündlicher Leistungen gibt es keine präzisen rechtlichen Vorgaben. Daher dürfen Lehrkräfte nach eigenem pädagogischem Ermessen entscheiden, wie sie diese Leistungen ermitteln. Sie dürfen auch „die individuelle Verschiedenheit der Schüler in Leistungsentwicklung und Lernverhalten berücksichtigen“ (vgl. ebd., S. 432).

¹ Sachfremde Bewertungsgegenstände sind z. B. persönliche Familienumstände des Lernenden oder die familiäre Unterstützung beim Lernen.

Der Lehrperson bleibt es außerdem selbst überlassen, wie sie mit der Dokumentation der Leistungen der mündlichen Mitarbeit verfahren möchte. Sie ist nicht rechtlich verpflichtet regelmäßig, z. B. täglich, Aufzeichnungen anzufertigen. Bei einer fehlenden freiwilligen Beteiligung am Unterricht eines Lernenden ist die Lehrkraft jedoch dazu verpflichtet, sich durch die direkte Ansprache oder Leistungsaufforderung einen Eindruck von dessen Leistungsstand und Leistungsbereitschaft zu verschaffen (vgl. Böhm 2008, S. 33 ff.).

Im Folgenden werden Auszüge aus dem zweiten Abschnitt „Leistungsbewertung“ des Hessischen Schulgesetzes dargestellt. Zunächst erfolgen allgemeine Vorgaben zur Leistungsbewertung in § 73:

„(1) Die Leistungen der Schülerinnen und Schüler werden durch Noten oder Punkte bewertet, soweit die Leistungen für die Erteilung von Zeugnissen und entsprechenden Nachweisen erheblich sind [...].

(2) Grundlage der Leistungsbeurteilung sind die mündlichen, schriftlichen, praktischen und sonstigen Leistungen, die die Schülerin oder der Schüler im Zusammenhang mit dem Unterricht erbracht hat. Für die Leistungsbewertung sind die im Unterricht vermittelten Kenntnisse und Fähigkeiten maßgebend.

(3) Zuständig für die Bewertung einzelner Schülerleistungen und für die Gesamtbewertung der im Beurteilungszeitraum erbrachten Leistungen sind auch bei inklusiver Beschulung die Lehrerinnen und Lehrer, die die Schülerin oder den Schüler in dem jeweiligen Fach zuletzt unterrichtet haben. Die Beurteilung des Arbeits- und Sozialverhaltens für den Beurteilungszeitraum erfolgt durch die Klassenkonferenz.

(4) Bei der Beurteilung durch Noten (Punkte) ist folgender Maßstab zugrunde zu legen:

1. sehr gut (15/14/13), wenn die Leistung den Anforderungen in besonderem Maße entspricht,

2. gut (12/11/10), wenn die Leistung den Anforderungen voll entspricht,

3. befriedigend (9/8/7), wenn die Leistung im Allgemeinen den Anforderungen entspricht,

4. ausreichend (6/5/4), wenn die Leistung zwar Mängel aufweist, aber im Ganzen den Anforderungen noch entspricht,

5. mangelhaft (3/2/1), wenn die Leistung den Anforderungen nicht entspricht, jedoch erkennen lässt, dass die notwendigen

Grundkenntnisse vorhanden sind und die Mängel in absehbarer Zeit behoben werden können,

6. ungenügend (0), wenn die Leistung den Anforderungen nicht entspricht und selbst die Grundkenntnisse so lückenhaft sind, dass die Mängel in absehbarer Zeit nicht behoben werden können.

Ist eine Leistungsbewertung aus Gründen, die die Schülerin oder der Schüler zu vertreten hat, nicht möglich, so erhält sie oder er die Note ungenügend (0)“ (HSchG § 73, S. 61).

Unter § 74 werden Vorgaben für Zeugnisse definiert:

„(1) Der Leistungsstand der Schülerinnen und Schüler wird unter Angabe der Leistungsbewertung in den einzelnen Fächern in Zeugnissen, schriftlichen Berichten oder in anderer, dem Bildungsgang entsprechender Form ausgewiesen.

(2) Ein allgemeines Zeugnis wird am Ende eines jeden Schuljahres oder Ausbildungsabschnittes oder als Übergangszeugnis beim Verlassen der Schule erteilt. Ein Halbjahreszeugnis am Ende des ersten Halbjahres eines Schuljahres informiert über den aktuellen Leistungsstand, das Zeugnis am Ende des Schuljahres weist den Leistungsstand aus, der während des Schuljahres erreicht wurde“ (HSchG § 74, S. 62).

Die Informationsrechte der Eltern und der SuS sind unter § 72 des Hessischen Schulgesetzes zu finden:

„(1) Die Schülerinnen und Schüler und ihre Eltern sind in allen wichtigen Schulangelegenheiten zu informieren und zu beraten. Dazu gehören [...] Grundzüge der Planung und Gestaltung des Unterrichts, Grundzüge der Unterrichtsinhalte und Unterrichtsziele sowie der Leistungsbewertung einschließlich Versetzung und Kurseinstufung [...].

(3) Die Schulleiterin oder der Schulleiter sowie die Lehrerinnen und Lehrer sollen die Eltern und Schülerinnen und Schüler in angemessenem Umfang informieren und beraten über [...] die Leistungsbewertung [...]“ (HSchG §72, S. 60).

Die Koordination der Leistungsbewertung ist die Aufgabe der Fach- und Fachbereichskonferenzen:

„Fach- und Fachbereichskonferenzen beraten über alle ein Fach, eine Fachrichtung oder einen Lernbereich betreffenden Angelegenheiten. Sie entscheiden im Rahmen der Rechts- und Verwaltungsvorschriften und

der von der Schul- oder Gesamtkonferenz beschlossenen Grundsätze insbesondere über [...] die Koordination der Leistungsbewertung“ (HSchG §134, S. 109).

In den Verordnungen zum Hessischen Schulgesetz werden unter § 19 allgemeine Grundsätze zur Leistungsfeststellung und –bewertung definiert:

„Die Leistungsfeststellung und Beurteilung nach § 73 des Hessischen Schulgesetzes [...] stützt sich auf die Beobachtungen im Unterricht und auf die mündlichen, schriftlichen und, sofern solche vorgesehen sind, die praktischen Leistungsnachweise und Leistungskontrollen. Leistungsfeststellung und -bewertung beziehen sich auf die gesamte Lernentwicklung der Schülerin oder des Schülers und umfassen sowohl die fachlichen Fähigkeiten, Kenntnisse und Fertigkeiten sowie die Leistungsbereitschaft, als auch Aussagen über das Verhalten der Schülerin oder des Schülers, wie es sich im Schulleben darstellt.

Hierbei ist zu beachten, dass Leistungsbewertung ein pädagogischer Prozess ist, der im Dienste der individuellen Leistungserziehung steht und der sich nicht nur auf das Ergebnis punktueller Leistungsfeststellungen, sondern auf den gesamten Verlauf der Lernentwicklung der Schülerin oder des Schülers bezieht. Der Verlauf der Lernentwicklung ist daher in die abschließende Leistungsbewertung einzubringen und soll der Schülerin oder dem Schüler eine ermutigende Perspektive für die weitere Entwicklung eröffnen“ (Hessisches Kultusministerium 2006, S. 18).

In Anlage 3 der Verordnungen zum Hessischen Schulgesetz sind folgende Hinweise zur Beurteilung des Arbeits- und Sozialverhaltens festgeschrieben:

„Nach § 2 des Hessischen Schulgesetzes soll die Schule den Schülerinnen und Schülern die dem Bildungs- und Erziehungsauftrag entsprechenden Kenntnisse, Fähigkeiten und Werthaltungen vermitteln. Die Schülerinnen und Schüler sollen insbesondere lernen, [...]

- sich für sich und andere einzusetzen sowie die Fähigkeiten zur Zusammenarbeit und zum sozialen Handeln zu entwickeln,
- Konflikte vernünftig und friedlich zu lösen, aber auch Konflikte zu ertragen [...].

Der Unterricht muss durch angemessene inhaltliche, didaktische und methodische Ansätze aber auch durch fach-, klassen-, jahrgangs- oder schulformübergreifenden Unterricht versuchen, diesen Anforderungen

gerecht zu werden und den Schülerinnen und Schülern den Erwerb überfachlicher Qualifikationen wie [...]

- Kooperationsbereitschaft und Kommunikationsfähigkeit,
- Organisation und Ausführung von Arbeitsaufträgen,
- Anwendung von Lern- und Arbeitstechniken,
- Gemeinschaftsfähigkeit, Fairness und Hilfsbereitschaft,
- Lernbereitschaft, Sorgfalt und Konzentrationsfähigkeit,

ermöglichen, denen auch in der Berufs- und Arbeitswelt zunehmende Bedeutung zukommt. Durch ermutigende Hinweise in den Lern- und Arbeitsprozessen entwickeln Schülerinnen und Schüler entsprechende Fähigkeiten. Lob und Anerkennung fördern den Erziehungsprozess und stärken ihre Persönlichkeit“ (Hessisches Kultusministerium 2006, S. 28).

Im Abschnitt 3 des § 25 der Verordnungen zum Hessischen Schulgesetz finden sich Anhaltspunkte für die Gewichtung der „mündliche Note“. In solchen Fächern, in denen zwei schriftlichen Klassenarbeiten pro Halbjahr geschrieben werden, bestimmt sie rund die Hälfte der Gesamtnote und in Fächern mit nur einer schriftlichen Klassenarbeit pro Halbjahr sogar zwei Drittel:

„In den Fächern [...] (Deutsch, Mathematik, 1., 2. und 3. Fremdsprache, Griechisch) machen die schriftlichen Arbeiten die Hälfte der Grundlagen der Leistungsbeurteilung aus, in den übrigen Fächern etwa ein Drittel“ (Hessisches Kultusministerium 2006, S. 21).

In diesen Verordnungen gibt es zudem unter § 23 Abs. 2 Vorgaben zu den Informationsrechten von Lernenden und ihren Eltern:

„Zu Beginn eines Schuljahres sollen die Schülerinnen und Schüler und die Eltern darüber informiert werden, nach welchen Gesichtspunkten die Bewertung ihrer Leistungen erfolgt. [...] Darüber hinaus sind Schülerinnen und Schüler mindestens einmal im Schulhalbjahr über ihren mündlichen Leistungsstand zu unterrichten“ (Hessisches Kultusministerium 2006, S. 20).

In der Oberstufen- und Abiturverordnung (OAVO) wird bezüglich der Bewertungsgegenstände für die „mündlichen Note“ folgendes festgelegt:

„Für die Bewertung der Leistungen am Ende eines Schulhalbjahres sind die im Unterricht kontinuierlich erbrachten Leistungen mindestens so bedeutsam wie die Ergebnisse der Leistungsnachweise. Zu den im

Unterricht kontinuierlich erbrachten Leistungen gehören vor allem die Mitarbeit im Unterricht, Versuchsbeschreibungen und -auswertungen, Protokolle, schriftliche Ausarbeitungen, Präsentationen, Hausaufgaben, Referate und solche schriftlichen Leistungen, welche die Schülerin oder der Schüler in Absprache mit der Lehrkraft des jeweiligen Kurses im Zusammenhang mit Unterrichtsinhalten auf eigenen Wunsch erbringt“ (OAVO § 9 (3), S. 9).

In den rechtlichen Vorgaben zur Leistungsbewertung werden also folgende Grundsätze deutlich:

In die Bewertung sollen nur die Leistungen einfließen, die im Unterricht von der Lehrperson beobachtet werden. Zudem muss die Leistung einem Lernenden generell eindeutig zugeordnet werden können und eine Gleichbehandlung gegeben sein. Prinzipiell hat die Lehrkraft einen Ermessensspielraum, der jedoch nicht mit Beliebigkeit verwechselt werden darf.

Nach dem Hessischen Schulgesetz soll Leistungsbewertung explizit nicht nur auf das Ergebnis punktueller Leistungsüberprüfungen bezogen werden, sondern auch den Verlauf der Lernentwicklung des Lernenden berücksichtigen. Die Abstufung der Noten (-punkte) ist dabei klar geregelt. Dagegen sind die Informationsrechte der SuS und Eltern nur vage formuliert („im angemessenen Umfang zu informieren“). In den Verordnungen ist hingegen präziser geregelt, dass zu Beginn jedes Schuljahres Informationen zu der Leistungsbewertung angekündigt werden und die SuS mindestens einmal im Schulhalbjahr Auskunft über ihren Leistungsstand erhalten müssen.

Die mündliche Note oder die mündliche Mitarbeit werden weder im Hessischen Schulgesetz noch in den Verordnungen explizit genannt. Die Gewichtung der mündlichen Note im Vergleich zu den schriftlichen Leistungen ergibt sich eher indirekt aus dem Abschnitt 3 des § 25 der Verordnungen zum Hessischen Schulgesetz. Die konkreten Bewertungsgegenstände für die mündliche Note lassen sich aus dem § 9 der Oberstufen- und Abiturverordnung herleiten. Dort wird die „Mitarbeit“ als einer der Bewertungsgegenstände erwähnt. Das Arbeits- und Sozialverhalten wird getrennt von den übrigen Leistungen beurteilt und ist daher kein Gegenstand der mündlichen Mitarbeit. Für die Kriterien der Bewertung mündlicher Mitarbeit gibt es keine rechtlichen Vorgaben. Sie liegen daher im Ermessen der Lehrperson.

3. Theoretischer Hintergrund

Ausgehend von den elementaren Leistungsbegriffen wird in diesem Kapitel zunächst definiert, was die mündliche Mitarbeit als Teil mündlicher Leistungen auszeichnet. Anschließend wird allgemein auf den Prozess der Leistungsbewertung inklusive der Gütekriterien und Bezugsnormen eingegangen. Auf dieser Basis wird sich mit der Bewertung mündlicher Mitarbeit auseinandergesetzt. Dabei werden mögliche Bewertungskriterien, Vorteile und Schwierigkeiten sowie vorliegende empirische Befunde über die Schulpraxis dargestellt. Zum Schluss erfolgen eine Zusammenfassung der theoretischen Grundlagen und ein Ausblick auf die Studie dieser Arbeit.

3.1 Leistungsbegriffe im schulischen Kontext

In diesem Abschnitt wird zuerst dargelegt, was unter Leistung verstanden werden kann und wie sich dies auf die Schule übertragen lässt. Hierbei wird auch auf das Leistungsprinzip im Zusammenhang mit Zeugnissen und Noten eingegangen. Anschließend erfolgt die Abgrenzung mündlicher von schriftlichen Leistungen. Dabei wird besonders darauf eingegangen, inwiefern die Begriffe „mündlichen Note“, mündliche Leistungen und mündliche Mitarbeit zusammenhängen.

3.1.1 Leistung

„Leistung ist das Ergebnis einer Untergruppe menschlicher Handlungen“ (Heckhausen 1974, S. 11). Dabei besteht eine Handlung in einem Verhalten aus anhaltenden, zielgerichteten Aktivitäten, die einen Anfang und ein Ende haben. Leistung ist also das Handlungsziel bestimmter Aktivitäten. Die Voraussetzungen für das Ausüben einer Handlung sind einerseits, dass die handelnde Person in der Lage und andererseits auch gewillt ist, die Handlung auszuführen. Das Können sowie die Erwartung auf Erfolg bzw. Misserfolg hängen von dem Verhältnis zwischen der Schwierigkeit der Handlung und der persönlichen Fähigkeit ab. Das Bemühen bzw. die Anstrengungsbereitschaft erfordert eine Intention des Handelnden. Um die Fähigkeit einer handelnden Person zu ermitteln, werden ihre Handlungsergebnisse mit denen anderer Personen in derselben Situation verglichen. Die Schwierigkeit des intendierten Handlungsziels ergibt sich daraus, wie viele ähnliche Personen diese Handlung verglichen mit anderen Handlungen erfolgreich abschließen.

Um die Untergruppe von Handlungssituationen genauer zu bestimmen, nennt Heckhausen fünf Bedingungen einer „leistungsthematischen Handlungssituation“, die zwar der psychologischen Sichtweise entspringen, jedoch als elementar angesehen werden können, um Leistung zu charakterisieren:

Erstens muss die Handlung zu einem Ergebnis führen, das objektivierbar ist und Aufgabencharakter aufweist.

Zweitens muss das Handlungsergebnis auf einen Maßstab hinsichtlich der Schwierigkeit und des benötigten und tatsächlich aufgebrauchten Kraftaufwandes bezogen und danach beurteilt werden können.

Drittens wird vorausgesetzt, dass die leistungsthematischen Handlungen gelingen und misslingen können. Dafür müssen die Anforderungen, die Schwierigkeiten und Kraftaufwand betreffen, in einem Intervall zwischen zu leicht und zu schwer liegen. Viertens müssen für die Handelnden subjektiv verbindliche Maßstäbe vorhanden sein und als diese anerkannt werden: Ein Gütemaßstab für einen bestimmten Schwierigkeitsgrad und einen Maßstab für den Kraftaufwand. Diese Maßstäbe werden als sogenannte Tüchtigkeitsmaßstäbe zusammengefasst. Die Beschäftigung mit diesen kennzeichnet leistungsmotiviertes Handeln.

Fünftens muss der Handelnde selbst der Verursacher des Handlungsziels sein. Bei unbeabsichtigten, zufälligen, von außen beeinflussten oder erzwungenen Handlungen wird die Verantwortung nicht beim Handelnden gesehen und so liegt keine leistungsthematische Situation vor.

Aus diesen Voraussetzungen des Leistungsbegriffs leitet Heckhausen zwei verschiedene Perspektiven her. Liegt ein Gütemaßstab vor, wird der Leistungsbegriff meist fähigkeitszentriert interpretiert, wobei eine leistungsthematische Handlung immer ein Minimum an Anstrengung erfordert. Als Alternative dazu kann Leistung anstrengungszentriert, d. h. anhand einer Anstrengungsskala über den Kraftaufwand beurteilt werden. Im Wesentlichen wirken beide Ausprägungen, die Fähigkeit als stabiles Merkmal und die Anstrengung als variables Attribut, nicht exklusiv, sondern miteinander. Ausschlaggebend für die Wahl der Perspektive ist vor allem die vom Beobachter abhängige Ursachenzuschreibung für eine erbrachte Leistung, die eher auf die Fähigkeit oder die Anstrengungsbereitschaft des Handelnden zurückgeführt wird (vgl. Heckhausen 1974, S. 11 ff.). Ein anstrengungsorientierter Leistungsbegriff

allein bietet besonders im schulischen Kontext, ein gewisses Konfliktpotential¹ (vgl. Heckhausen 1974, S. 60).

In der Physik spielt der Kraftaufwand bei der Definition von Leistung eine große Rolle. Leistung (P) wird als Arbeit (W) pro Zeiteinheit (t) verstanden. Dabei ist die Arbeit definiert als Kraft (F) multipliziert mit dem zurückgelegten Weg (s) (vgl. Stolz 2005, S. 26 ff.).

$$P = \frac{dW}{dt}, \text{ wobei } W = F \cdot s$$

Die Herkunft des Begriffs „Leistung“ kann u. a. in den romanischen Sprachen entdeckt werden. Das lateinische Verb „praestare“ (vorstehen, sich auszeichnen) wird aus „sta“ hergeleitet, was z. B. „Stand“ oder „Stelle“ bedeutet und heute noch im italienischen Wort für Leistung „prestazione“ erkennbar ist.

Im Gegensatz zum romanischen, eher statischen Leistungsbegriff geht der Ursprung von Leistung in der indogermanischen Sprache eher mit einer prozesshaften Bedeutung einher. Der Begriff der Leistung kann auf das indogermanische Wort „leis“ zurückgeführt werden, was „Fußspur“ oder „Furche“ bedeutet. Es wird bestimmt durch das Wort „lis“ für „gehen“ Außerdem bedeutet das gotische Wort „laistjan“ „eine Spur nachverfolgen“ (vgl. Ziegenspeck 1999, S. 30).

Daraus wird die normative Bedeutung abgeleitet, dass das Erbringen von Leistungen dem Folgen einer Spur bzw. dem Vorgeben und dem Orientieren an Anforderungen und Gütemaßstäben entspricht. Dementsprechend geben Leistungen von SuS einen Einblick darüber, inwiefern sie mit ihrer Handlung den Anforderungen und Maßstäben gerecht werden (vgl. Sacher 2014, S. 13).

3.1.2 Schulleistung

„Unter Schulleistung versteht man zusammengefasst die von der Schule initiierten Lernprozesse und Lernergebnisse der Schüler“ (Ingenkamp & Lissmann 2008, S. 131). Daher kann der Begriff der Schulleistung auch synonym mit Lernleistung verwendet werden (vgl. Jäger 2007, S. 45). Im schulischen Kontext wird Leistung in vierfacher Bedeutung verwendet: Schulleistung ist zuerst eine Forderung seitens der Schule, denn sie wird von den Lernenden abverlangt. Der damit verbundene

¹ Bei gleichem Leistungsergebnis und unterschiedlichem Kraftaufwand zweier Lernenden wird es womöglich nicht als gerecht aufgefasst, wenn diese unterschiedlich gewertet werden.

Druck hat auch eine motivationale Funktion¹. Ferner wird die Schulleistung als generelle Tätigkeit des Lernenden gesehen, bei der die Leistungserbringung im Lernen und sonstigen Arbeiten besteht, welche Fleiß und Ausdauer erfordern. Als spezifische Schulleistung gelten die einzelnen domänenspezifischen Tätigkeiten und die damit verbundenen unterschiedlich gut ausgeprägten Leistungen. Schließlich wird Schulleistung als Beitrag für Andere verstanden. Dabei besteht die Vorstellung, dass die Schule die Lernenden für die Fort- und Weiterbildung qualifiziert und damit indirekt einen Beitrag nicht nur für das Individuum, sondern für die Gesellschaft leistet. Durch den Aufbau von Wissen und Kompetenzen profitieren insbesondere spätere Institutionen, wie die Universität oder der Arbeitgeber (vgl. Furck 1967, S. 118 ff.).

Aufgrund verschiedenster sprachlicher Ursprünge zur Bestimmung des Leistungsbegriffs, die entweder dynamische Prozesse oder statische Ergebnisse beschreiben, weist dieser eine hohe Komplexität auf. Außerdem wird er je nach Kontext unterschiedlich verstanden und es kann daher keine einheitliche Definition für die pädagogische Praxis geben.

Schulleistung wird oft als Ergebnis einer Anstrengung (s. o.) definiert. Daher wird meist der produktorientierte Leistungsbegriff angewandt. Wird Leistung dagegen prozessorientiert gesehen, steht nicht das Produkt² im Vordergrund, sondern der Vorgang der Anstrengung, der zum Ergebnis geführt hat (vgl. Beckmann 1978, S. 44 ff.).

Schulleistung ist nach Krapp definiert als „das Ergebnis von Lernprozessen, die durch Unterrichtsmaßnahmen angeregt und/ oder planvoll gesteuert wurden“ (Krapp 1984, S. 46). Hierbei lassen sich unter dem Begriff der Schulleistung mehrere Abstraktionsebenen unterscheiden. Relativ konkret und wahrnehmbar ist die empirische Ebene von Schulleistung als die jeweils von einem Lernenden erbrachte Einzelleistung (z. B. Ergebnis in einem Test). Auf einer abstrakteren Ebene wird Schulleistung gemäß eines hypothetischen Konstrukts als überdauernde, situations- und zeitübergreifende Fähigkeit aufgefasst. Daher wird eine Differenzierung in das direkt beobachtbare „Schulleistungsverhalten“ und die nicht direkt beobachtbare „Schulleistungsfähigkeit“ vorgeschlagen. Die quantitativ gewonnenen Erkenntnisse über das Schulleistungsverhalten eines Lernenden

¹ Hierbei handelt es sich um Fremdkontrolle und damit um extrinsische Motivation. Für eine detaillierte Beschreibung siehe Heckhausen (1980, S.459).

² Das Produkt bzw. das Ergebnis einer Anstrengung ist im schulischen Kontext z. B. ein Aufsatz, ein Plakat oder das Ergebnis einer Klassenarbeit.

(durch Noten, Zeugnisse o. Ä.) werden als Indikatoren für die Schulleistungsfähigkeit herangezogen (vgl. Krapp 1984, S. 46 f.).

Nach Heller (1984, S. 15 ff.) wird unter der Schulleistung eines Lernenden „das gesamte Leistungsverhalten im Kontext schulischer Bildungsbemühungen“¹ verstanden. Für die Leistung soll dabei der Lernprozess genauso beachtet werden wie das Lernprodukt. Auch Klafki fordert eine Relativierung und Erweiterung des produktorientierten Leistungsbegriffs (vgl. Klafki 2007, S. 75 f.). Er versteht unter Leistung im schulischen Kontext das Ergebnis und den Vollzug² einer zielgerichteten Tätigkeit, welche mit Anstrengungen und ggf. Selbstüberwindung einhergehen und für die Gütemaßstäbe gelten. Um neben Leistungsergebnissen auch kognitive Prozesse im Unterricht (wie z. B. kommunizieren, argumentieren und kritisieren) bewerten zu können, braucht es geeignete Leistungskriterien, die jedoch schwer zu objektivieren sind (vgl. Klafki 2007, S. 228 f.).

3.1.3 Leistungsprinzip

Um das Leistungsprinzip in der Schule darzustellen, erfordert es zunächst die Definition des gesellschaftlichen Leistungsprinzips. „Das Leistungsprinzip ist ein gesellschaftliches Verteilungsprinzip, welches Chancen und Positionen nach Leistung zuweist – nicht etwa nach Abstammung, Bekanntschaft, weltanschaulicher Nähe usw.“ (Sacher 2014, S. 15). Konkurrierende Prinzipien sind das Vorrecht der Geburt, das Anciennitätsprinzip, Ideologieprinzip, Beliebtheitsprinzip und das Sozialprinzip. Die Verteilung erfolgt hierbei (analog zur obigen Reihenfolge) nach der sozialen Herkunft, nach dem Alter oder der Dauer der Zugehörigkeit, nach der Angehörigkeit zu einer Weltanschauung oder nach dem Bekanntheits- oder Beliebtheitsgrad. Nach dem Sozialprinzip der Gesellschaft erhalten auch jene Unterstützung, die unverschuldeter Weise keine eigenen Leistungen erbringen können, z. B. durch Krankheit, Alter oder Behinderung.

Da das Leistungsprinzip zwar dominiert aber nicht allein besteht, sollte die deutsche Gesellschaft lediglich als „leistungsorientierte Gesellschaft“ bezeichnet werden (vgl. Sacher 2014, S. 15). Das Leistungsprinzip ist im Sprachgebrauch häufig positiv

¹ Dabei wird nicht nur das Schülerverhalten, sondern auch das der Eltern und der Lehrperson miteinbezogen. Auf die zahlreichen Bedingungsfaktoren von Schulleistung und Lernerfolg wird in diesem Kontext nicht näher eingegangen. Siehe dazu das Angebots-Nutzungs-Modell von Helmke (2014, S. 71), Heller (1984, S.54 ff.) und für Determinanten von Schulleistung Jäger (2007, S.48 ff.).

² Dies stellt einen Unterschied zu der Definition von Heckhausen und zu der von Krapp (s. o.) dar, in der lediglich das Ergebnis der Tätigkeit als Leistung betrachtet wird.

konnotiert, da es mit Betriebsamkeit und Fleiß als Tugenden assoziiert wird (vgl. Heckhausen 1974, S. 58).

Voraussetzung für eine adäquate Umsetzung des Leistungsprinzips in der Schule sind eindeutige und anerkannte Maßstäbe zur Erfassung der individuellen Leistungen sowie eine Chancengleichheit aller beteiligten Akteure¹ (vgl. Klafki 2007, S. 221 ff.). „Wie jede Wertvorstellung enthält Chancengleichheit eine Forderung, die im Grunde nicht nur einfach, sondern zugleich auch evident ist, d. h. sie bedarf weder einer Begründung noch läßt sie sich in Frage stellen oder gar leugnen. Es ist die Evidenz des Gleichheitsgrundsatzes. Keiner soll bei der gesellschaftlichen Zuweisung von Chancen, aus Gründen, die er nicht selbst zu verantworten hat, gegenüber anderen bevorzugt oder benachteiligt werden. So einleuchtend und simpel dieser Grundsatz ist, so problematisch und schwierig wird es, wenn man ihn auf konkrete und komplexe Sachverhalte anzuwenden versucht“ (Heckhausen 1974, S. 107). Es soll nun ein kurzer Abriss über die Entstehung des Leistungsprinzips in der Schule folgen:

Im Mittelalter waren der Stand und die damit verbundenen Chancen auf gesellschaftliche Positionen allein abhängig von der Geburt in die jeweiligen Lebensverhältnisse. Seit dem 17. und 18. Jahrhundert bildete sich vor dem Hintergrund des an Bedeutung gewinnenden Bürgertums und der Anforderungen der zentral verwalteten absolutistischen Staaten ein staatliches Schulsystem heraus. Aufgrund der wirtschaftlichen Entwicklungen und neuen Anforderungen bezüglich der Effektivität brauchte es ein zweckmäßigeres Verteilungsprinzip, gerade im Hinblick auf Beruf- und Lebenschancen. Außerdem griff das Leistungsprinzip die Idee des Gleichheitsanspruchs der Französischen Revolution auf und wurde zu einem emanzipatorischen Element für das Bürgertum (vgl. Nipkow 1978, S. 7 ff.). Schließlich wurde im 19. Jahrhundert der Schulabschluss schrittweise Voraussetzung für bestimmte Berufsausbildungen und gesellschaftliche Positionen. Zunächst galt dies nur für gehobene Tätigkeiten, die den Abschluss eines Gymnasiums erforderten. Dieser Prozess dehnte sich dann aber schließlich im 20. Jahrhundert auf eine große Zahl der Berufe aus, für die vorher keine Real- oder Gymnasialbildung notwendig war. So wurde das Leistungsprinzip zum Bindeglied zwischen Schule und Gesellschaft (vgl. Klafki 2007, S. 211 f.).

Dennoch wird das Leistungsprinzip in der Pädagogischen Psychologie teilweise

¹ Davon ausgenommen werden hier Menschen mit besonders schweren körperlichen oder geistigen Einschränkungen.

scharf kritisiert. „Das Leistungsproblem erweist sich als ein grundlegendes Entscheidungsproblem der Erziehungspraxis außerhalb und innerhalb des institutionalisierten Bildungswesens und folglich als eine Kernfrage der Erziehungstheorie“ (Klafki 2007, S. 209). Einerseits wird die Bezeichnung der Gesellschaft als „Leistungsgesellschaft“ und die daraus resultierende „Leistungsschule“ angesichts anderer bestehender Verteilungsprinzipien angezweifelt.

Andererseits gibt es aus pädagogischer Perspektive Zweifel am Leistungsprinzip: Zum einen wird kritisiert, dass Lernende vor dem Eintritt in das Arbeitsleben eine gewisse Schonzeit vor dem Leistungsdruck bräuchten, um sich als Persönlichkeiten zu entwickeln. Zum anderen ist die Chancengleichheit als notwendige Voraussetzung (s. o.) in der Schule nicht gewährleistet¹. Der schulische Wettbewerb um Lebenschancen ist heute noch sehr stark von der sozialen Herkunft abhängig² (vgl. Sacher 2009, S. 17 ff.).

3.1.4 Zeugnisse und Noten

Parallel zur Entwicklung des Leistungsprinzips gab es einen wachsenden Bedarf nach Zeugnissen und Zensuren. Nach Ziegenspeck ist ein Zeugnis ursprünglich als eine Aussage eines Zeugen über einen Sachverhalt zu verstehen, der beurteilt und geprüft wird. Das Wort „Zensur“ leitet sich aus dem lateinischen Wort „censere“ ab, was so viel bedeutet wie begutachten oder einen Wert beimessen. Im schulischen Kontext fasst ein Zeugnis verschiedene Leistungen eines Lernenden zusammen, die durch Zensuren bzw. Noten beziffert werden (vgl. Ziegenspeck 1976 S. 34). Ein Schulzeugnis ist die „schriftlich fixierte, schulamtliche Zeugenaussage über den Schüler für einen interessierten Dritten. (Es) [...] ist das zusammenfassende Ergebnis von Zensuren, d. h. von klassifizierten Beurteilungen und Bewertungen von Leistungen und Verhaltensweisen. Der Lehrer ist in diesem Zusammenhang Urteilender und Zeuge zugleich“ (Dohse 1967, S. 6).

Die Zensur wird als Ziffernote ausgedrückt. Die Noten ermöglichten ab dem 16. Jahrhundert die Zuweisung eines Rangplatzes zu einem Schüler/ einer Schülerin innerhalb der Klassenrangfolge. Diese Rangfolge nach Leistung bestimmte in diesen Zeiten sogar die Sitzordnung der Lernenden. Die geschichtliche Entwicklung

¹ Auf die Kritik am Leistungsprinzip in der Schule kann in dieser Arbeit nicht näher eingegangen werden.

² Siehe dazu Rolff (1997, S.19 ff.) und für die aktuelle, wenig veränderte Situation die Pisa Ergebnisse in OECD (2016, S. 61 ff.).

der Zensuren als Leistungsstufen verlief jedoch nicht geradlinig oder für alle Regionen Deutschlands einheitlich (vgl. Dohse 1967, S. 46 ff.). Es waren verschiedenste Skalen und Notenstufen (und auch verbale Beurteilungen) verbreitet, bis sich die heute noch geltende sechsteilige Abstufung¹ in der Zeit des Nationalsozialismus etablierte und nach dem Beschluss der Kultusminister 1954 für die Bundesrepublik Deutschland übernommen wurde (vgl. Ziegenspeck 1976, S. 34 ff.). Heute wird diese Notenstufung für die deutsche gymnasiale Oberstufe durch ein Punktesystem im Intervall von 0-15 Punkten ersetzt (vgl. OAVO Hessen 2014, § 9).

Als Vorläufer der heutigen Zeugnisse gelten die sogenannten Benefizienzeugnisse, die für bedürftige Lernende im 16. Jahrhundert als Empfehlungsschreiben für ein Stipendiat dienten. Erst im 18. und 19. Jahrhundert entwickelten sich flächendeckend Schulzeugnisse, zunächst in Form von meist verbalen Beurteilungen. 1788 wurde in Preußen das Reifezeugnis als Abschlusszeugnis des Gymnasiums (der heutigen Allgemeinen Hochschulreife) eingeführt, welches unter dem Aspekt des allgemeinbildenden Charakters die Würdigkeit für ein Studium bescheinigte. Allerdings war dieser Abschluss nicht zwingend erforderlich für ein Universitätsstudium. Erst Mitte des 19. Jahrhunderts zeichnete sich das Reifezeugnis durch den „allgemeinen Berechtigungscharakter“ aus, den es heute noch aufweist (vgl. Ziegenspeck 1976, S. 34 ff.).

Welche Funktionen erfüllen Zeugnisse und Noten? Walter Dohse formulierte folgende Funktionen: Kontrollfunktion, rechtliche Funktion, Anreizfunktion, Zuchtfunktion und Orientierungsfunktion (vgl. Dohse 1967, S. 62 ff.). Diese wurden von Ziegenspeck aufgegriffen und zusammengefasst zur Orientierungs- und Berichtsfunktion, die pädagogische Funktion und die Auslese-, Rangierungs- und Berechtigungsfunktion (vgl. Ziegenspeck 1976 S. 51 ff).

Sacher (2014, S. 22 ff.) unterscheidet folgende Funktionen und klassifiziert sie zudem in zwei Kategorien: Einerseits gesellschaftliche, politische und schulorganisatorische Funktionen, zu der Selektion und Stigmatisierung, Sozialisation, Legitimation, Kontrolle und Prognose gezählt werden und andererseits unterrichtlich und erzieherische Funktionen, wie Prognose, Information und Rückmeldung, Disziplinierung und Lehr- und Lerndiagnose.

Dabei wird die Menge an Funktionen von dem Autor als „überfrachtet“ kritisiert und eine Fokussierung auf unterrichtliche und erzieherische Funktionen gefordert.

¹ „Sehr gut (1), gut (2), befriedigend (3), ausreichend (4), mangelhaft (5), ungenügend (6)“

Der Umstieg von freien Formulierungen auf Ziffernnoten schaffte eine Verbesserung der Vergleichbarkeit und erleichterte somit vor allem die Selektion (vgl. Sacher 2014, S. 21). Dies kann jedoch zu Konkurrenzdenken führen und wird nicht zuletzt deshalb kontrovers diskutiert. Als Alternative zur Notengebung werden daher zum Teil verbale Beurteilungen gefordert, die wiederum auch kritisiert werden, da sie nur eine andere Darstellungsform des Urteils der Lehrperson sind und sich für den Beurteilungsprozess keine Veränderung ergibt. Befürworter der verbalen Beurteilungen weisen allerdings darauf hin, dass sich eine Lehrkraft durch das Schreiben solch einer Beurteilung zwangsläufig differenzierter mit den Leistungen des Lernenden auseinandersetzen muss (vgl. Valtin 2002, S. 11 ff.). Dieser Diskurs kann in dieser Arbeit nicht in allen Einzelheiten ausgeführt werden, sollte aber zumindest eine Erwähnung finden.

3.1.5 Mündliche Leistungen

In diesem Abschnitt werden, ausgehend von einem Modell der Kommunikation, die Unterschiede von Mündlichkeit und Schriftlichkeit aufgezeigt sowie die Facetten mündlicher (Schul-)Leistungen dargestellt. Ferner wird auf die Kompetenzfelder mündlicher Leistungen eingegangen und der Begriff der „mündliche Note“ von den mündlichen Leistungen abgegrenzt.

Nach dem psychologischen Modell der menschlichen Kommunikation¹ von Schulz von Thun (2008, S. 25 ff.) übermittelt der Sender dem Empfänger eine Nachricht, wie in Abbildung 2 dargestellt wird. Diese hat vier Seiten: Eine Nachricht vermittelt nicht nur einen bestimmten Sachinhalt, also Daten und Fakten, sondern besitzt noch mehr Ebenen:

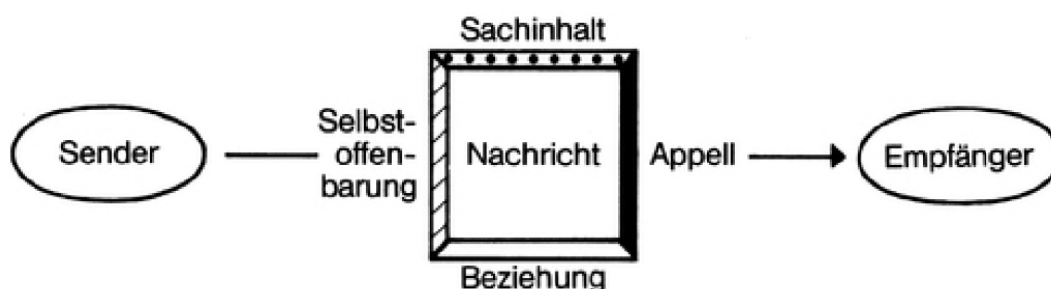


Abbildung 2: Die vier Seiten (Aspekte) einer Nachricht nach Schulz von Thun (2008, S. 30)

Jede Nachricht enthält in unterschiedlich starker Ausprägung Informationen über den Sender, der durch die Äußerung, teilweise unbewusst, etwas über sich selbst

¹ Dieses Modell vereint den Inhalts- und Beziehungsaspekt von Watzlawick (2011, S. 61 ff.) mit dem Organon-Modell der Sprachfunktionen von Bühler (1999, S. 24 ff.).

offenbart. Auf der Beziehungsebene einer Nachricht wird deutlich, wie Sender und Empfänger zueinander stehen. Durch bestimmte Formulierungen, Körpersprache, Tonfall usw. kann der Sender Wertschätzung, Verachtung, Bevormundung etc. zeigen, die vom Empfänger unterschiedlich gedeutet werden können. Zudem haben fast alle Nachrichten eine Funktion bzw. sollen eine bestimmte Wirkung beim Empfänger erzielen, die entweder offen oder verdeckt artikuliert wird.

Nach Fiehler (2009, S. 25 ff.) ist jede Form der Verständigung ein Ausdruck von Kommunikation. Die verschiedenen kommunikativen Praktiken erfordern die Berücksichtigung eines Gefüges von sozialen Regeln und Konventionen und können entweder schriftlicher oder mündlicher Art sein. Beispiele für Ersteres sind Briefe, E-Mails oder Aufsätze und für das Zweite Gespräche oder Redebeiträge. Mündliche Kommunikation ist ein Prozess, an dem verschiedene Sinne gleichzeitig beteiligt sind. Denn sie ist multimodal, d. h. sie findet parallel auf verschiedenen Ebenen statt. Verbale mündliche Kommunikation vereint alles mit dem Mund Gesprochene. Körperliche Kommunikation ist ein Sammelbegriff für Mimik, Gestik, Körperhaltung und Körperkonstellationen. Hinzu kommt die aufgrund von visuellen Wahrnehmungen und Schlüssen vollzogene Kommunikation.

Im Vergleich zu schriftlicher Kommunikation ist die mündliche Kommunikation eine Handlung, die leicht umgesetzt werden kann. Hierfür wird meist weniger Aufmerksamkeit und Konzentration aufgewendet als bei geschriebener Sprache. Das gesprochene Wort ist flüchtig und stark kontextbezogen, das Geschriebene hingegen wenig kontextbezogen und von Dauer. Dadurch ergeben sich unterschiedliche Funktionen von Mündlichkeit und Schriftlichkeit:

„Mündliche Kommunikation hat ihre zentrale Funktionalität im Bereich *der interaktiven Bewältigung aktueller Situationen*. Zentrale Funktionen sind dabei die unmittelbare wechselseitige Beeinflussung und Steuerung und die Vermittlung von Wissen. Geschriebene Sprache hingegen [...] hat ihre spezifische Funktion in der *raum-zeitlichen Distribution und Tradierung von Texten*. Sie ist damit das zentrale Instrument der Wissensvermittlung“ (Fiehler 2009, S. 27, Hervorhebung im Original).

Damit mündliche Kommunikation gelingen kann, müssen einige Voraussetzungen erfüllt sein. Mindestens zwei Menschen verständigen sich zweckgebunden. Dabei sind sie einander zugänglich, d. h. sie können miteinander in Kontakt treten und haben einen gemeinsamen Bezugsraum. Sie können sich gegenseitig wahrnehmen und nehmen räumliche und mentale Perspektiven ein. Die Verständigung ist flüchtig, erfolgt multimodal und die einzelnen Nachrichten bedingen sich

gegenseitig. Diese folgen zeitlich aufeinander und können in folgende verschiedene Einheiten aufgeteilt werden: Laut, Wort, funktionale Einheit, Gesprächsbeitrag und Gespräch (vgl. Fiehler 2009, S. 29 ff.).

„Kommunikation ist zentrales Medium von Unterricht“ (Nuding 2006, S. 14). Mündliche Leistungen in der Schule sind alle auf mündlicher Kommunikation beruhenden Unterrichtsleistungen der Lernenden (vgl. Arnold & Jürgens 2001, S. 76). Diese mündlichen Unterrichtsleistungen, die in allen Schulfächern vorkommen können, sind beispielsweise Vorträge, das mündliche Vorstellen einer Lösung oder Ähnlichem, Referate, Präsentationen und die mündliche Mitarbeit¹ (vgl. Krumwiede u. a. 2014, S. 23). Diese sind mitunter schwer von schriftlichen Leistungen zu trennen, da z. B. bei einem Referat eine schriftliche Vorarbeit erfolgt, die dann noch mitunter von dem Lernenden auswendig gelernt werden kann (vgl. Koch & Oesterreicher 2007, S. 349).

Alle mündlichen Leistungen erfordern bestimmte Fähigkeiten und Fertigkeiten seitens der SuS. Dazu hat Ulf Abraham, ausgehend von den Vorgaben zum „Sprechen und Zuhören“ der Bildungsstandards der KMK² für den Deutschunterricht, ein Kompetenzmodell für mündliche Leistungen entwickelt. Die Überschneidungen der einzelnen Kompetenzen werden in Abbildung 3 dargestellt.

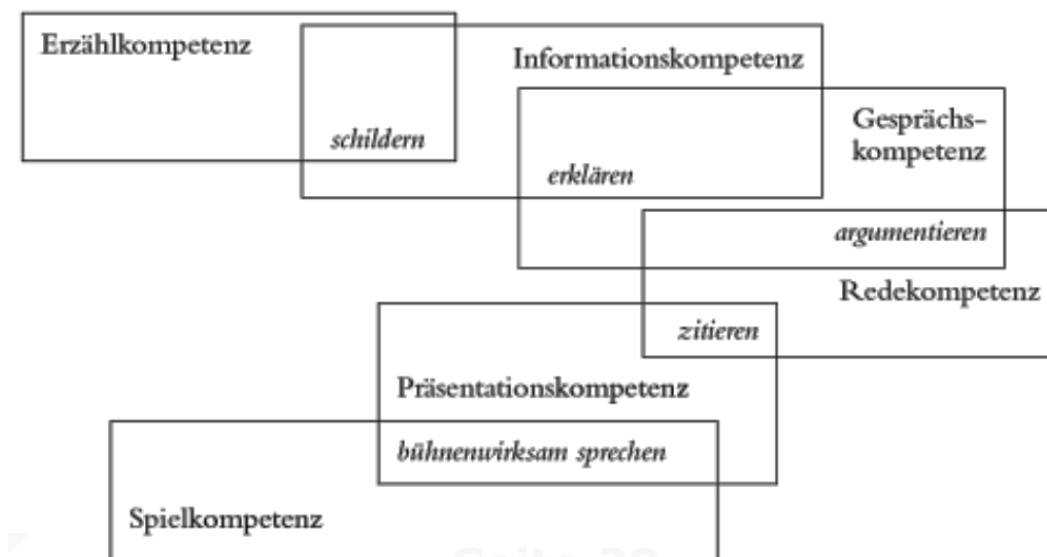


Abbildung 3: Kompetenzen der Mündlichkeit nach Abraham (2008, S. 29)

Lernende sollen in der Lage sein, sich an einem Gespräch zu beteiligen, eigene Argumente einzubringen oder die Moderation zu übernehmen. Dafür erfordert es

¹ Auf den wichtigen Begriff der „mündlichen Mitarbeit“ wird im nächsten Abschnitt ausführlich eingegangen.

² Für weitere Informationen siehe Kultusministerkonferenz (KMK) (2012, S. 15).

eine *Gesprächskompetenz*. Um einen Sachverhalt anderen berichten, zusammenfassen oder erklären zu können benötigen Lernende die *Informationskompetenz*. Die *Erzählkompetenz* beinhaltet, dass SuS Geschichten für den jeweiligen Hörer nachvollziehbar und spannend erzählen können. Bei Vorführungen oder Rollenspielen sollen Lernende sich in eine Rolle hineinversetzen und diese überzeugend verkörpern können (Spielkompetenz). Sie benötigen *Redekompetenz*, wenn sie vor Publikum einen Vortrag oder ein Referat halten oder in einer Debatte einen passenden Beitrag leisten. Wenn sie eigene oder fremde Texte zitieren oder vortragen bedarf es einer gewissen *Präsentationskompetenz* (vgl. Abraham 2008, S. 29 ff.).

Für die mündliche Note¹ dienen alle mündlichen Unterrichtsleistungen der SuS als Bewertungsgegenstände. Hinzu kommen aber auch schriftliche, praktische und sonstige (fachspezifischen) Leistungen, die im Unterricht erbracht werden². Dazu zählen u. a. kleinere Lernkontrollen, wie Vokabeltests oder Vorführungen vor der Klasse (Anschrieb an der Tafel), die Erledigung und Qualität von Hausaufgaben, die Heftführung sowie das Vorhandensein der nötigen Arbeitsmaterialien (vgl. Arnold & Jürgens 2001, S. 76).

Die mündliche Mitarbeit nimmt den größten Anteil der mündlichen Unterrichtsleistungen ein (vgl. Nickel 2004, S. 2). Sie soll als ein für alle Fächer und Jahrgangstufen fester Bestandteil³ der mündlichen Schulleistung im nächsten Abschnitt näher beleuchtet werden.

3.1.6 Mündliche Mitarbeit

Die mündliche Mitarbeit eines Lernenden beschreibt die spontanen mündlichen Beiträge zum Unterrichtsgespräch⁴ (vgl. Göller 1968, S. 81; Sacher 2014, S. 157). Unterrichtsgespräche sind eine auf das schulische Lehren, Lernen und Erziehen abzielende mündliche Kommunikation im Unterricht⁵. Anders als bei einem

¹ Hierfür wird auch oft der Begriff „sonstige Mitarbeit“ verwendet (vgl. Souvignier & Förster 2014; Disselkamp 2004).

² In der mündlichen Note werden also, bis auf die schriftlichen Klassenarbeiten, sämtliche (auch schriftliche) Leistungen der Lernenden zusammengefasst. Daher besteht Kritik an der Begrifflichkeit der mündlichen Note. Siehe dazu Krieger (2003).

³ Die übrigen mündlichen Leistungen werden dagegen je nach Fach und Lehrperson variabel eingefordert.

⁴ Vorbereite Beiträge (Referate, Vorstellen der Hausaufgaben), die zum Teil schriftlich vorher ausgearbeitet wurden, werden nicht dazu gezählt.

⁵ Das Lernen findet also in einem interaktiven Gesprächsprozess statt. Diese Prozesshaftigkeit fällt unter den erweiterten Leistungsbegriff nach Heller (1984) und Klafki (2007), wie in 3.1.2 *Schulleistung* dargelegt wurde.

Unterrichtsgespräch zwischen SuS besteht bei jenen zwischen Lehrperson und SuS ein asymmetrisches Rollen-Verhältnis. Unterrichtsgespräche finden meist auf Initiative der Lehrkraft statt und sind auch aufgrund der schulischen Inhalte von einem lockeren Alltagsgespräch zu differenzieren (vgl. Bittner 2006, S. 20 ff.). Unter Unterrichtsgesprächen werden sowohl Großgruppengespräche¹ (mit der gesamten Klasse) als auch Gespräche in kleineren Gruppen, z. B. bei Gruppenarbeiten oder Partnerarbeiten, verstanden (vgl. Ritz-Fröhlich 1982, S. 22 ff.).

Nach Birkel (1984, S. 233 f.) kann eine Lehrperson aus den Beiträgen der SuS zum Unterrichtsgespräch auf das Gelingen ihres Unterrichts schließen: An den Äußerungen der Lernenden kann sie erkennen, inwieweit die Lernenden mit den Unterrichtsinhalten und den Lernmethoden zurechtkommen und diese dann ggf. verändern, um sie an die SuS anpassen.

„Die lebendige Gestaltung von Unterricht kann überhaupt nicht auf die mündliche Mitarbeit der Schüler verzichten“ (Birkel 1984, S. 234).

Die aktive Mitarbeit im Unterricht wird daher von LuL sehr geschätzt, wobei allein die Bereitschaft, sich am Unterrichtsgeschehen zu beteiligen, schon mit Fleiß assoziiert wird (vgl. Sacher 2014, S. 157). Der „Schülerjob“ besteht darin, am Unterricht aktiv teilzunehmen. Durch eine Meldegeste zeigt ein Lernender in einem Großgruppengespräch ein Angebot an, einen Sprachbeitrag zu leisten. Dabei ist es, nach einer weit verbreiteten Haltung der Lernenden unerheblich, ob dieser Lernende auch für das Sprechen ausgewählt („drangenommen“) wird. Die aufgezeigte Bereitschaft wird als ein Zeichen des Engagements wahrgenommen und gewertet. Es zählt also nicht nur die objektive Leistung des Lernenden, sondern auch das subjektive Bemühen. Zudem sinkt die Wahrscheinlichkeit einer, von SuS ungewollten, Aufforderung durch die Lehrperson, etwas zum Gespräch beizutragen (vgl. Breidenstein 2006, S. 98 ff.).

In den rechtlichen Grundlagen des Landes Hessen findet sich die Forderung nach Betriebsamkeit z. B. unter den allgemeinen Grundsätzen zur Leistungsbeurteilung² wieder. Demnach sollen für die Leistungsfeststellung und -bewertung nicht nur die fachlichen Fähigkeiten, Kenntnisse und Fertigkeiten, sondern auch die Leistungsbereitschaft sowie das Verhalten des Lernenden berücksichtigt werden. Dies beschränkt sich nicht nur auf Hessen. In der Notenverordnung des Landes

¹ Im Gegensatz hierzu wird für gewöhnlich bei den sogenannten „Sozialformen“ im Unterricht das Unterrichtsgespräch als Großgruppengespräch von einer Partnerarbeit oder Gruppenarbeit abgegrenzt.

² Siehe dazu § 19 der Verordnungen zum Hessischen Schulgesetz in *Kapitel 2 Rechtliche Vorgaben*.

Baden Württemberg wird z. B. festgeschrieben: „Mitarbeit bezieht sich vor allem auf den *Arbeitswillen*, der sich in Beiträgen zu den selbständig oder gemeinsam mit anderen zu lösenden Aufgaben äußert“ (NVO 2016, § 6 (2), eigene Hervorhebung). Aus den Voraussetzungen mündlicher Kommunikation (s. o.) können folgende Aussagen über die Mitarbeit im Unterricht abgeleitet werden:

Die mündliche Mitarbeit der Lernenden ist abhängig von der Gruppen- bzw. Klassengröße. Je größer die Gruppe des stattfindenden Gesprächs ist, desto mehr Zuschauer gibt es und desto größer ist die Hemmschwelle, sich zu beteiligen. Mündliche Kommunikation erfolgt nicht ohne eine andauernde, gegenseitig wertende Wahrnehmung der beteiligten Akteure. Daher erfordert die mündliche Mitarbeit gewisse verbale Fähigkeiten, wie z. B. Konflikt- und Argumentationsfähigkeit (vgl. Krumwiede u. a. 2014, S. 18 ff.). Lernende benötigen daher für die mündlichen Beiträge zum Unterrichtsgespräch eine grundlegende Gesprächsfähigkeit¹ (vgl. Potthoff u. a. 1996, S. 14 f.). Auf die Gesprächsfähigkeit wird hier, als Kompetenz der Mündlichkeit und für die mündliche Mitarbeit bedeutsame Voraussetzung, näher eingegangen.

„Gesprächsfähig sind Kinder, wenn sie sachgemessen und partnerbezogen miteinander sprechen. Sie finden sich in verschiedenen gegebenen Gesprächssituationen zurecht und können Sprecher- und Hörrollen einnehmen, einen Gegenstand oder einen Sachverhalt richtig erfassen und verständlich wiedergeben, als Person authentisch sein und sich in der Kommunikation und Zusammenarbeit mit anderen an vereinbarte Regeln halten. In schwierigen Situationen ergreifen sie die Möglichkeiten zur Metakommunikation und thematisieren problematische Punkte“ (Potthoff u. a. 1996, S. 14).

Daraus ergeben sich für die Gesprächsfähigkeit die drei Teilkompetenzen zuhören, argumentieren und moderieren:

Das *Zuhören* ist im Gespräch eine kognitive und interaktive Herausforderung des Verstehens. Ferner erfordert ein Gespräch die Kompetenz, in Aushandlungen und beim Problemlösen *argumentieren* zu können und kohärente Beiträge zu leisten sowie jene Gespräche *moderieren* zu können (vgl. Berkemeier & Spiegel 2014, S. 123 ff.).

¹ Die Begriffe Gesprächskompetenz (s. o.) und Gesprächsfähigkeit werden synonym verwendet. Siehe dazu z. B. Becker-Mrotzek & Brünner (2004, S. 81) und Berkemeier & Spiegel (2014, S. 120 ff.).

3.2 Leistungen feststellen, beurteilen und bewerten

Im letzten Abschnitt wurden die Leistungen im Bereich Schule dargestellt, insbesondere die der mündlichen Mitarbeit. Diese Leistungen müssen von den Lehrkräften festgestellt, beurteilt bzw. bewertet werden. Für den gesamten Prozess von der erbrachten Leistung bis zur Notengebung gibt es keine einheitlichen Begriffe:

Die Pädagogische Diagnostik beschreibt alle Tätigkeiten, welche die Voraussetzungen für Lehr- und Lernprozesse bestimmen, die Lernprozesse analysieren und Lernergebnisse feststellen. Mithilfe diagnostischer Instrumente soll ein Verhalten beschrieben, die Gründe für dieses Verhalten erklärt und damit auf zukünftiges Verhalten geschlossen werden (vgl. Ingenkamp & Lissmann 2008, S. 13).

Die pädagogische Diagnostik besteht aus zwei elementaren Prozessen¹, der „Leistungsmessung“ und „Leistungsbeurteilung“ (vgl. Jürgens 2010, S. 44; Mietzel 2007, S. 420; Paradies u. a. 2014; Ziegenspeck 1999, S. 129), oder auch bezeichnet als „Leistungsfeststellung“ und „Leistungsbeurteilung“ (vgl. Heckhausen 1974, S. 25 ff.; Heller 1984, S. 29; Hessisches Kultusministerium 2006, S. 18; Holmeier 2013, S. 108; Klauer 2014, S. 103; Neuweg 2006, S. 7), wobei die Begriffe „Beurteilen“ und „Bewerten“ meist bedeutungsgleich verwendet werden. Unter der ersten Phase, der Leistungsfeststellung bzw. -messung wird das Ermitteln der Schülerleistung, z. B. unter Verwendung eines Messinstruments, verstanden. Die Phase der Leistungsbeurteilung bzw. -bewertung beschreibt die Einordnung und Bewertung der bereits erfassten Schülerleistungen anhand bestimmter Vergleichsmaßstäbe². Die Autoren verzichten dabei teilweise gänzlich auf den Gebrauch eines übergeordneten Begriffs für den gesamten Prozess oder verwenden dafür einen der oben genannten Begriffe (vgl. ebd.; Winter 2016, S. 3 ff.; Sacher 2014, S. 33 ff.).

Im Gegensatz dazu unterscheidet Jäger die Begriffe Beurteilung und Bewertung: „Das Beurteilen bezieht sich auf einen [...] Prozess, an dessen Anfang eine Beobachtung und an dessen Ende eine Beurteilung (Urteil) steht [...]. Bewertung bezeichnet das Zumessen eines Wertes oder einer Bedeutung zu einem Beobachtungsgegenstand“ (Jäger 2007, S. 32).

¹ Wobei die Übergänge fließend sind.

² Auf die genaue Bedeutung wird in diesem Abschnitt noch näher eingegangen.

Brünung und Saum (2009) schlagen ebenfalls eine differenziertere Unterscheidung¹ der Begrifflichkeiten vor. Hierbei wird der gesamte Prozess „Leistungsbewertung“ genannt, der aus drei Phasen besteht: Feststellen, beurteilen und benoten bzw. Leistungsmessung, Leistungsbeurteilung und Notengebung (vgl. Brüning & Saum 2009, S. 122 f.). Die Begriffe „Bewerten“ und „Benoten“ haben im schulischen Kontext dieselbe Bedeutung (vgl. Sacher 2014, S. 83).

In dieser Arbeit soll die überwiegend vorkommende Zweiteilung in Leistungsfeststellung- und -beurteilung zwar berücksichtigt werden, jedoch nicht ohne auf die Verschiedenheit der Vorgänge der Beurteilung und Bewertung einzugehen. Auch im alltäglichen Sprachgebrauch verschmelzen die Begriffe oft zu einem². Es gibt zudem keinen allgemein gültigen Oberbegriff für den gesamten Prozess. Aus Gründen der besseren Lesbarkeit wird daher in dieser Arbeit mit (Leistungs-)Bewertung (als Ziel des Ganzen) bzw. Bewertungspraxis, bis auf die explizite Trennung in diesem Abschnitt, der gesamte Prozess bezeichnet.

3.2.1 Leistungsfeststellung

In der Phase der Leistungsfeststellung wird ermittelt, welche Leistungen die Lernenden erbracht haben und welche nicht (vgl. Klauer 2014, S. 103). Die einzelnen Methoden der Schulleistungsdiagnostik werden für gewöhnlich in „objektive“, d. h. messende Verfahren bzw. Tests, und „subjektive Verfahren“, wie z. B. Lehrerurteile, gegliedert³ (vgl. Heller 1984, S. 28).

Daher soll sich in diesem Abschnitt zum einen mit dem Messen und zum anderen mit subjektiven Prüfverfahren befasst und sie auf die messtheoretischen Gütekriterien bezogen werden.

3.2.1.1 Objektive Verfahren

Die klassische Testtheorie bietet die Basis für standardisierte Schulleistungstests (vgl. Heller 1984, S. 24). Diese können entweder formell, wie z. B. allgemeine Schulleistungs- oder Intelligenztests, oder informell sein, wie Lernkontrolltests (vgl.

¹ Bei der Leistungsmessung werden Leistungen beobachtet, gemessen oder erhoben. Bei der Leistungsbeurteilung werden Stärken und Schwächen der Schülerleistung identifiziert. Zum Schluss folgt ggf. die Notengebung, bei der der Schülerleistung eine Ziffernote zuordnet wird.

² Dies wird auch in *Kapitel 2 Rechtliche Vorgaben* deutlich.

³ Die Bezeichnungen bedeuten jedoch nicht, dass Messverfahren per se objektiv und Lehrerurteile niemals objektiv sind.

ebd., S. 143 ff.). Dabei dienen (meist von Experten) entwickelte Messinstrumente zur Erfassung von Schulleistungen (vgl. Heller & Hany 2014, S. 91 f.).

Stevens (1966, S. 21 ff.) versteht unter Messen das Zuordnen von Zahlen zu Objekten nach bestimmten Regeln. Dabei sind die Merkmale eines Objektes der Gegenstand der Messung und nicht das Objekt selbst. Von ihrer Ausprägung wird dann auf das Objekt geschlossen. Das Ergebnis einer Messung schließt noch keine Bewertung mit ein.

Diese Vorstellung wird auch als Messen im engeren Sinne bezeichnet, welches auf drei verschiedenen Skalenniveaus geschehen kann: Ordinalskalenniveau (zur Festlegung einer Rangfolge), Intervallskalenniveau (Aussagen über Abstände) und das Verhältnisskalenniveau für Aussagen über Größenverhältnisse (vgl. Sacher 2014, S. 33 f.).

3.2.1.2 Gütekriterien

Jede wissenschaftliche Messmethode muss im Sinne der Qualitätssicherung bestimmten Gütekriterien genügen. Die drei wichtigsten Hauptgütekriterien sind die *Objektivität*, *Reliabilität* und *Validität*.

Nach Sacher (2014, S. 35 ff.) sind Messungen dann *objektiv*, wenn sie unabhängig vom Messenden erfolgen. Weiter lässt sich die Objektivität in drei Phasen differenzieren: Erstens die Objektivität bei der Durchführung der Messung in Bezug auf Regeln und Einheitlichkeit. Zweitens meint die Auswertungsobjektivität inwiefern gleiches Verhalten einer Testperson stets auf die gleiche Weise ausgewertet wird. Die Interpretationsobjektivität bezeichnet den Grad, inwiefern gleiche Testwerte von verschiedenen Personen auf dieselbe Weise interpretiert werden.

Die *Reliabilität* bezeichnet die Zuverlässigkeit einer Messung, d. h., dass die Ergebnisse genau und nicht durch Messfehler verfälscht sind. Um dies zu überprüfen, gibt es drei verschiedene Möglichkeiten. Durch eine Wiederholung desselben Tests, die Aufteilung des Tests in zwei Hälften mit getrennter Auswertung sowie die Möglichkeit der Paralleltestmethode, bei der zwei Varianten eines Tests getrennt durchgeführt werden.

Die *Validität* einer Messung bezeichnet die Gültigkeit, also inwiefern der Test auch wirklich das Merkmal misst, welches gemessen werden soll.

Zusätzlich zu den Hauptgütekriterien gibt es Nebengütekriterien, die hier noch kurz erwähnt werden sollen: Ist das Verfahren *ökonomisch*, so bleibt der Materialverbrauch und Zeitaufwand für die Durchführung in einem vernünftigen Maß. Die Messung sollte außerdem eine *Nützlichkeit* aufweisen und außerdem

zumutbar sein. Zuletzt soll eine *Akzeptanz*, durch die Übereinstimmung von Ergebnissen und Meinungen bzw. Wertmustern, gegeben sein (vgl. Sacher 2014, S. 35 ff.). Die Gütekriterien definieren damit einen hohen Anspruch an „objektive Verfahren“ der Leistungsfeststellung.

3.2.1.3 Subjektive Verfahren

Zu den „subjektiven Verfahren“ zählen alle nicht-messenden Methoden, um Schulleistungen zu ermitteln. Darunter fallen alle konventionellen, schulischen Methoden wie schriftliche Prüfungen (Klassenarbeiten, Schulaufgaben) und mündliche Prüfungen sowie die (wissenschaftliche) Verhaltensbeobachtung und Ratingverfahren. Diese Methoden besitzen in Bezug auf die wissenschaftlichen Gütekriterien eine geringere Objektivität, Zuverlässigkeit und Gültigkeit als „objektive Verfahren“ (vgl. Heller 1984, S. 33).

Die Gewährleistung der Objektivität verlangt nach Sacher (2014, S. 46 f.) eine Vereinheitlichung und Schematisierung der Verfahren, was die Bandbreite legitimer Leistungen und damit die Kreativität einschränken kann. Objektivität kann im Schulalltag nur in wenigen, geregelten Situationen und am ehesten noch bei schriftlichen Prüfungen gewährleistet sein. Wenn die Objektivität als grundlegendes Gütekriterium angestrebt wird, entsteht die Gefahr der zu starken Fokussierung auf z. B. schriftliche Leistungen, die zu Lasten der Lernprozesse und Beobachtungen geht. Daraus ergibt sich folgendes Dilemma in der pädagogischen Praxis: „Das Bemühen um die Sicherung von Objektivität geht somit ersichtlich auf Kosten der Validität: Das, was einigermaßen objektiv erhoben und beurteilt werden kann, repräsentiert die Schülerleistung [...] nur noch sehr ausschnitthaft und unvollständig“ (Sacher 2014, S. 47).

Die Herausforderung bei der schulischen Leistungsfeststellung besteht allgemein darin, dass die Lehrperson durch ihren Unterricht unmittelbar am Geschehen der Schulleistungen beteiligt ist. Dazu legt sie in einem großen Maße selbst fest, welche Leistungen für die Feststellung relevant sind und wie die Leistungen ermittelt werden¹. Zudem liegt die Durchführung allein bei der Lehrperson (vgl. Sacher 2014, S. 83; Ziegenspeck 1999, S. 130).

Die Leistungsfeststellung geschieht innerhalb zwischenmenschlicher Beziehungen. Schulleistungen können weiterhin nur ab bestimmten „Schwellen“ sichtbar werden. Wie nah Lernende tatsächlich an der geforderten Leistung oder bestimmten

¹ Beispielsweise gestalten Lehrkräfte ihre Klassenarbeiten selbst.

Schwellen liegen, ist oft für die Lehrperson nicht erkennbar¹. Zusätzlich können die Leistungen der SuS nur stichprobenartig und somit nicht vollständig ermittelt werden (vgl. Sacher 2014, S. 104 f.).

Besonders bei mündlichen Leistungen spielen Beobachtungen der Lehrkräfte eine zentrale Rolle für die Leistungsfeststellung (vgl. Krumwiede u. a. 2014, S. 19). Lernprozessbegleitendes Beobachten ist nach Nuding (2006, S. 7 f.) eine der zentralen Aufgaben einer Lehrperson und bildet die Basis für die spätere Beurteilung. Eine Beobachtung ist eine absichtliche, planvolle und selektive Form der Wahrnehmung (vgl. ebd., S. 30 f.). „Wir sprechen von Beobachtung, wenn aus einem Ablauf von Ereignissen etwas aktiv, also nicht beiläufig, zum Objekt der eigenen Aufmerksamkeit gemacht wird“ (Bortz 1984, S. 191).

„Ziel einer Verhaltensbeobachtung als diagnostischer Methode ist es, [...] offenes oder verdecktes Verhalten einer Person oder mehrerer Personen sowie deren Interaktionen [...] deskriptiv zu erfassen“ (Bodenmann 2006, S. 151).

Hierbei werden einige Formen² von Beobachtungsmethoden differenziert: Unsystematische „Gelegenheitsbeobachtungen“ erfolgen ohne bestimmte Kategorien und dienen dazu, sich einen ersten Eindruck zu verschaffen. Dagegen gibt es bei systematischen Beobachtungen klare Regeln zur Durchführung und Auswertung. Diese Beobachtungen können sich entweder auf das eigene Verhalten (Selbstbeobachtung) oder das anderer Personen (Fremdbeobachtung) beziehen. Bei einer Fremdbeobachtung kann der Beobachter Teil des Beobachtungskontextes sein (teilnehmende Beobachtung) oder von außen beobachten (nicht-teilnehmende Beobachtung). Eine Feldbeobachtung ist eine Beobachtung im natürlichen Umfeld des Geschehens, wohingegen eine Laborbeobachtung unter künstlichen, aber besser kontrollierbaren Bedingungen stattfindet (vgl. Bodemann 2006, S. 151 ff.).

Bei der Beobachtung von Schülerverhalten im Unterricht handelt es sich meist um teilnehmende Fremd- und Feldbeobachtungen, die eher unsystematisch erfolgen (vgl. Langhorst 1984, S. 211 f.). Da schulische Situationen sehr komplex sind und ein häufiger, fließender Wechsel zwischen aktiver Teilnahme und teilnehmender Beobachtung vollzogen wird, kann eine Lehrperson das Unterrichtsgeschehen nicht

¹ Dieses Problem wird auch Schwellenproblem genannt.

² Die hier genannten Beobachtungsmethoden stellen eine Auswahl dar und beanspruchen somit keine Vollständigkeit.

gänzlich wahrnehmen. Beobachtungen können ein Verhalten bzw. Situation ohnehin niemals vollständig abbilden (vgl. Boer & Reh 2012, S. 70).

Jede Beobachtung wird durch die persönliche Interpretation des Beobachtungsgegenstandes beeinflusst. Daher können dieselben Situationen von Beobachtenden mit unterschiedlichen Wert- oder Harmonievorstellungen verschieden beobachtet werden. Zudem spielen kognitive Faktoren, wie Konzentration, Logik und Erinnerungsvermögen eines Beobachtenden eine große Rolle (vgl. Jäger 2007, S. 64). Hinzu kommen außerdem persönliche Vorlieben, soziale Vorurteile und implizite Persönlichkeitstheorien. Insofern ist es für Menschen eine große Herausforderung, das Erscheinungs- und Verhaltensbild anderer objektiv zu beobachten¹. Daher ist es empfehlenswert, Beobachtungen z. B. mit Gedächtnisprotokollen zu dokumentieren (vgl. Langhorst 1984, S. 208 ff.). Sacher (2014, S. 157) empfiehlt, gezielt einzelne Lernende systematisch in Standardsituationen zu beobachten, um eine gewisse Gleichmäßigkeit zu gewährleisten. Allerdings können unter Umständen auch wiederholte, beiläufige Gelegenheitswahrnehmungen von bestimmten Verhaltensweisen einen, sogar objektiven, Gesamteindruck bilden (vgl. Langhorst 1984, S. 214).

Der Grund für die tendenziell geringere Erfüllung der Gütekriterien liegt also an den grundlegenden Unterschieden zwischen „objektiven“ und „subjektiven“ Verfahren zur Leistungsfeststellung. Dies relativiert zwar die Übertragbarkeit der Gütekriterien auf die Verfahren der schulischen Praxis, sie behalten jedoch weiterhin ihre Berechtigung (vgl. Wengert 2011, S. 329 ff.). Die Gütekriterien gelten dementsprechend auch für schulische Leistungsfeststellung als Leitlinie und deren Einhaltung als erstrebenswert (vgl. Tulodziecki u. a. 2010, S. 24 f.).

3.2.2 Leistungsbeurteilung und -bewertung

Bei der Leistungsbeurteilung und -bewertung werden die bereits ermittelten Leistungen, im Blick auf bestimmte Kriterien, eingestuft (vgl. Klauer 2014, S. 103). Hierbei wird zuerst mithilfe der Relation zu Normen und Erwartungen eine Aussage über die Qualität der Leistung getroffen. Danach kann dieser Qualität ein bestimmter Wert beigemessen werden. Die eigentliche Leistungsbeurteilung erfordert nicht zwingend eine Bewertung im Sinne einer Note. Eine Note basiert jedoch immer auf der vorherigen Beobachtung und Beurteilung (vgl. Brüning & Saum 2009, S. 123).

¹ Da die Beobachtung eine Form der visuellen Wahrnehmung ist, werden einige der Beobachtungsschwierigkeiten auch in der Wahrnehmungspsychologie thematisiert (vgl. Bortz 1984, S. 191).

3.2.2.1 Bezugsnormen

Bei der Beurteilung von Schülerleistungen müssen diese in Bezug zu einer Norm gesetzt werden, um zu bestimmen, welche als gut und welche als weniger gut eingestuft werden. Als Referenzrahmen werden im Allgemeinen drei Bezugsnormen unterschieden:

Bei der *sozialen Norm* werden die Leistungen eines einzelnen Lernenden im Vergleich mit der durchschnittlichen Leistung der gesamten Lerngruppe beurteilt.

Im Gegensatz dazu findet die Beurteilung nach *kriterialer Bezugsnorm* nach festen Kenntnis- bzw. Fähigkeitsmaßstäben, also unabhängig von der Gruppenleistung, statt. Maßgeblich ist der Vergleich mit dem vorgegebenen Lernziel.

Bei der *individuellen Norm* ist der individuelle Lernfortschritt des Lernenden entscheidend für die Beurteilung. Die persönlichen Verbesserungen oder Verschlechterungen werden bei dieser Norm damit besonders berücksichtigt (vgl. Rheinberg 2014, S. 59 ff.).

Oft wird bei der Anwendung der sozialen Bezugsnorm die Häufigkeit der vorkommenden Leistungen anhand der Normalverteilung (auch Gauß'sche Glockenkurve) ausgerichtet (vgl. Sacher 2014, S. 119). Durch soziale Vergleiche zwischen den SuS wird zudem ein „Big-Fish-Little-Pond“-Effekt begünstigt. Dabei ist die Ausprägung des akademischen Selbstkonzeptes¹ abhängig von den Vergleichen mit dem Können der anderen SuS in der Lerngruppe. In einer eher starken Lerngruppe hat dieser Lernende demnach ein geringeres Selbstkonzept als in einer schwächeren (vgl. Marsh & Parker 1984, S. 221 ff.). Das Selbstkonzept ist bedeutend für die Motivation und zukünftige Leistungen² (vgl. Helmke & van Aken 1995, S. 629 ff.). Dementsprechend wird die, in der Praxis häufig gebrauchte, soziale Bezugsnorm, im Hinblick auf den pädagogischen Nutzen kritisiert (vgl. Sacher 2014, S. 119 f.; Rheinberg 2014, S. 63 f.; Wengert 2011, S. 328).

Die Orientierung an der individuellen Bezugsnorm kann sich positiv auf die Motivation der Lernenden auswirken. Sie wirft allerdings die Frage auf, wie mit Unterschieden in den Verbesserungsmöglichkeiten verschieden starker Lernenden oder der Stagnation einer Leistung adäquat umgegangen werden kann (vgl. Sacher 2014, S. 122).

¹ Unter dem akademischen Selbstkonzept werden nach Möller & Köller (2004, S. 19) „generalisierte fachspezifische Fähigkeitseinschätzungen [verstanden], die Schüler und Studenten aufgrund von Kompetenzerfahrungen in Schul- bzw. Studienfächern erwerben“.

² Es besteht eine Wechselwirkung zwischen Selbstkonzept und Leistung.

Daher wird in der Literatur eine stärkere Betonung der kriterialen (und individuellen) Bezugsnorm für die schulische Praxis empfohlen (vgl. Rheinberg S. 69 ff.; Sacher 2014, S. 96; Tulodziecki u. a. 2010, S. 20).

3.2.2.2 Herausforderungen beim Beurteilen und Bewerten

Hat eine Lehrperson sich auf einen oder mehrere Bezugsrahmen festgelegt, können ihr noch Fehler beim Urteil unterlaufen. Die Gefahr, dass persönliche Vorlieben und Abneigungen die Beurteilung verfärben, lässt sich in sozialen Kontexten nie ganz vermeiden. Sacher (2014, S. 47 ff.) unterscheidet zwei Kategorien der Urteilsfehler: Zum einen die „ungleichmäßige Ausschöpfung des Beurteilungsspielraums“ und zum anderen „Interferenzen im Urteil (Voreingenommenheiten)“:

Zu Ersterem zählen *Strenge- bzw. Mildefehler*, bei denen eine Lehrperson entweder dazu neigt, kleine Mängel zu stark zu gewichten und damit wenig gute Urteile zu fällen oder Bewertungen verstärkt am oberen Ende des Spektrums zu vergeben. Bei der *Tendenz zur Mitte* werden Urteile am äußeren Spektrum vermieden, wohingegen bei der *Tendenz zu Extremurteilen* wenige Noten aus dem mittleren Spektrum gegeben werden.

Zu den Urteilsfehlern, die auf Voreingenommenheiten basieren, gehört der *Reihungsfehler*. Dabei beeinflusst die vorherige Leistung die Einschätzung für die nachfolgende (auch bei verschiedenen Lernenden). Bei einem *logischen Fehler* werden fälschlicherweise voreilige Schlüsse eines Leistungsmerkmals eines Lernenden auf ein völlig anderes geschlossen. Als *Halo-Effekt* wird eine Voreingenommenheit, z. B. über das Auftreten oder die Kleidung eines Lernenden, bezeichnet, die als globaler Gesamteindruck das Urteil über andere Merkmale, positiv oder negativ, überstrahlt (vgl. Sacher 2014, S. 47 ff.). Ergänzend sei noch der *Rosenthal-Effekt* genannt, bei dem Erwartungen und Einschätzungen der Lehrperson als „selbsterfüllende Prophezeiung“ den Urteilsprozess indirekt beeinflussen (vgl. Rosenthal & Jacobson 1966, S. 115 ff.).

Um den potentiellen Urteilsfehlern entgegenzuwirken, ist es zwingend erforderlich, dass Lehrkräfte ihre eigene Urteilspraxis reflektieren. Dafür ist es hilfreich, sich über die persönliche Urteilstendenz sowie bestimmte fachspezifische Traditionen¹ bewusst zu werden. Die Beobachtung bzw. Beschreibung der Leistung sollte zudem getrennt von der Beurteilung erfolgen. Ein Gesamturteil sollte möglichst spät, nach einer Vielzahl von beobachteten Leistungen, erfolgen und ggf. wieder revidiert

¹ Naturwissenschaften gelten eher als „strengere“ Fächer als z. B. musische und künstlerische Fächer.

werden. Außerdem können Reihungen, z. B. beim Korrigieren, aufgelockert und variiert werden. Insgesamt gilt die Maxime, im Sinne der Vermeidung unkontrollierter Wahrnehmung und Urteile, dass sich eine Lehrperson sowohl einen Eindruck vom Lernenden selbst als auch von dessen Leistung verschaffen soll (vgl. Sacher 2014, S. 48 ff.).

Neben der gesellschaftlichen, politischen und schulorganisatorischen Funktion der Leistungsbewertung, im Hinblick auf Laufbahntscheidungen, ist es für die Entwicklung der Lernenden bedeutsam, der unterrichtlich-erzieherischen Funktion gerecht zu werden, indem sie durch die Rückmeldung von Stärken und Schwächen sowie der Bereitstellung individueller Förderung in ihrem Lernprozess begleitet werden (vgl. Souvignier & Förster 2014, S. 9).

In den rechtlichen Vorgaben ist festgelegt, dass eine Lehrperson bestimmte Maßstäbe für die Leistungsbewertung¹ anlegen soll, jedoch immer einen gewissen pädagogischen Spielraum hat (siehe *Kapitel 2*). Auch aufgrund mangelnder Berücksichtigung von Gütekriterien und Urteilsfehlern kann aus der möglichen Verletzung dieser Ansprüche nicht zwangsläufig auf Fehler in der Leistungsbewertung geschlossen werden (vgl. Sacher 2014, S. 184).

Da jedoch die Beobachtung und Beurteilung oft in einer Note mündet und die Notengebung über grundsätzliche Lebenschancen der SuS entscheidet (vgl. Brüning & Saum 2009, S. 124) sollten Lehrpersonen, nicht nur die SuS, sondern auch sich selbst systematisch beobachten, um bestimmten Urteilstendenzen und Fehlern vorzubeugen (vgl. Langhorst 1984, S. 225).

Nützlich kann auch der Vergleich mit den Einschätzungen der SuS sein. Die Selbstbeobachtung bzw. -bewertung sowie die wechselseitige Bewertung werden daher als sinnvolle Ergänzung zur Fremdbeurteilung und -bewertung der Lehrperson gesehen (vgl. Winter 2016, S. 249 ff.).

¹ Hier ist der gesamte Prozess vom Beobachten bis zum Bewerten gemeint.

3.3 Mündliche Mitarbeit feststellen, beurteilen und bewerten

Mündliche Mitarbeit ist, neben den schriftlichen Leistungskontrollen, ein zentraler Bestandteil schulischer Leistungsbewertung (vgl. Brüning & Saum 2009, S. 122). Obwohl die mündliche Mitarbeit und andere mündliche Leistungen einen großen Teil des Unterrichtsalltags einnehmen, wird die Bewertung dieser Leistungen in der Literatur stark vernachlässigt oder gar nicht behandelt (vgl. Kirk 2004, S. 27; Souvignier & Förster 2014, S. 7). Wenn Mündlichkeit thematisiert wird, dann liegt meist ein Fokus auf mündlichen Prüfungen, wie z. B. in Birkel (1984, S. 229 ff.) oder Sacher (2014, S. 153 ff.).

In diesem Abschnitt werden zuerst mögliche Kriterien erarbeitet, die sich aus Abschnitt 3.1.6 *Mündliche Mitarbeit* ergeben, anhand derer die mündliche Mitarbeit bewertet werden kann. Im Anschluss daran werden zunächst Chancen der Bewertung mündlicher Mitarbeit, aber auch Grenzen diskutiert. Dabei sollen die im vorigen Abschnitt beschriebenen Probleme der Beobachtung und Beurteilung auf die mündliche Mitarbeit übertragen werden. Zudem wird es zwei Exkurse zur Berücksichtigung der Schülerpersönlichkeit und der Trennung von Lern- und Prüfungssituation geben. Zum Schluss werden Erkenntnisse aus den bereits durchgeführten Studien zu diesem Themenbereich dargestellt.

3.3.1 Kriterien

Aus den Vorgaben des Hessischen Schulgesetzes¹ lassen sich zunächst allgemein folgende Maßstäbe für die Bewertung mündlicher Mitarbeit ableiten: Für die Bewertung sind nur die im Unterricht vermittelten Kenntnisse und Fähigkeiten heranzuziehen. Hierbei sollen insbesondere nicht nur Leistungsergebnisse, sondern der gesamte Verlauf der Lernentwicklung der Lernenden berücksichtigt werden. Die Kriterien zur Bewertung mündlicher Mitarbeit² liegen im Ermessen der Lehrperson, sollen den SuS und Eltern jedoch zu Beginn des Schuljahres transparent gemacht werden. Die Lehrperson soll sich durch die Einforderung mündlicher Beiträge einen Eindruck von dem Leistungsstand eines Lernenden verschaffen (vgl. Böhm 2008, S. 33 ff.). Mündliche Leistungen müssen unabhängig von den schriftlichen Leistungen bewertet werden. Bei der Bewertung mündlicher

¹ Siehe dazu *Kapitel 2 Rechtliche Vorgaben*.

² Unter Kriterien werden in dieser Arbeit die relevanten Bewertungsgegenstände verstanden.

Mitarbeit sollte daher darauf geachtet werden, dass echte mündliche und keine indirekt schriftlichen Leistungen beurteilt werden (vgl. Wengert 2011, S. 342).

Im Mittelpunkt der Suche nach geeigneten Kriterien steht die schwierige Frage nach der Gewichtung von *Qualität und Quantität* der mündlichen Beiträge. Bei der Quantität spielt zum einen die *Freiwilligkeit* eine große Rolle, also ob Beiträge eher aufgrund eigener Initiative oder durch Nachfragen der Lehrperson zustande gekommen sind (vgl. Paradies u. a. 2014, S. 69). Zum anderen wird, neben der Freiwilligkeit, auch die Frequenz der Unterrichtsbeiträge, also die Regelmäßigkeit der wiederholten Beiträge, die *Kontinuität*, bewertet (vgl. Arnold & Jürgens 2001, S. 76; Disselkamp 2004, S. 53). Der Maßstab der Quantität resultiert zum einen aus der Notwendigkeit auf viele anregende Sprachbeiträge im Unterricht (vgl. Birkel 1984, S. 234) und zum anderen aus der Assoziation mit Fleiß bzw. Betriebsamkeit, was im Unterrichtsalltag als erwünscht gilt (vgl. Breidenstein 2006, S. 98 ff.). Während das Merkmal der Qualität der Unterrichtsbeiträge in der Literatur unstrittig ist, gilt das nicht für die Quantität:

Die mündliche Note soll nach Birkel nicht etwas zur Häufigkeit der Äußerungen aussagen, sondern nur zum Stand der Fähigkeiten und Fertigkeiten des Lernenden, also der Qualität (vgl. Birkel 1984, S. 234). Die Quantität ist demnach nur hilfreich, um einen Einblick in die Qualität der Beiträge und somit in den Lernstand der SuS zu erhalten, was Lehrkräfte notfalls durch gezieltes Aufrufen ebenfalls erreichen können (vgl. Wengert 2011, S. 342 ff.).

Zudem hat eine Untersuchung von Cohen, Rey und Sixtil (1969) gezeigt, dass die Häufigkeit und Intensität eines Merkmals bei diagnostischen Urteilsprozessen nicht getrennt voneinander beobachtet werden können, sondern zu einem Gesamteindruck verschmelzen. Daher sollte „eine gewisse Zurückhaltung beim Mündlichen, wenn sie auch ein Mangel ist, keinesfalls mit Nichtkönnen verwechsel[t]“ werden (Göller 1968, S. 81).

Neben der fachlichen Qualität und der Quantität der Beiträge kann es sinnvoll sein, die Gesprächsfähigkeit bzw. –kompetenz zu beurteilen, indem Merkmale der Teilkompetenzen *zuhören, argumentieren und moderieren* gezielt beobachtet werden (vgl. Berkemeier & Spiegel 2014, S. 123 ff.).

Potthoff, Steck-Lüschow & Zitzke schlagen dafür Kategorien vor, die in Tabelle 1 dargestellt werden.

Tabelle 1: Kategorien für die Beobachtung von Gesprächsfähigkeit nach Potthoff u. a. (1996, S. 14)

Kategorie	Merkmale
Was ist zu sehen?	Blickkontakt, Gestik, Haltung, Atmung
Was ist zu hören?	Artikulation, Melodieführung, Tempo/ Pausen/ Sprechfluss, Lautstärke/ Dynamik, Klangfarbe/ Stimme
Wie kann die Sprache beschrieben werden?	Wortwahl, Satzbau
Wie ist der Gesprächsbeitrag inhaltlich aufgebaut?	Themenbezug, Informationsmenge, Strukturierung des Gesprächs
In welcher Weise fließt die Gefühlsebene mit ein?	Meinung vertreten, Interesse zeigen, Ich-Botschaften, Authentizität
Wie stellt sich das Kind auf den Partner oder die Zuhörer ein?	Aufforderungen formulieren, Argumente finden, Akzeptanz und Kompromisse in der Kooperation

Zudem können für die Bewertung mündlicher Mitarbeit auch weitere Kompetenzen des Mündlichen¹ bedeutsam sein: Die Erzählkompetenz, Informationskompetenz, Redekompetenz und Spielkompetenz.

Die Kriterien zur Bewertung mündlicher Mitarbeit können gemeinsam mit den Lernenden erarbeitet werden (vgl. Krumwiede u. a. 2014, S. 24). Dazu sei die Mindmap in Abbildung 4 beispielhaft angeführt:

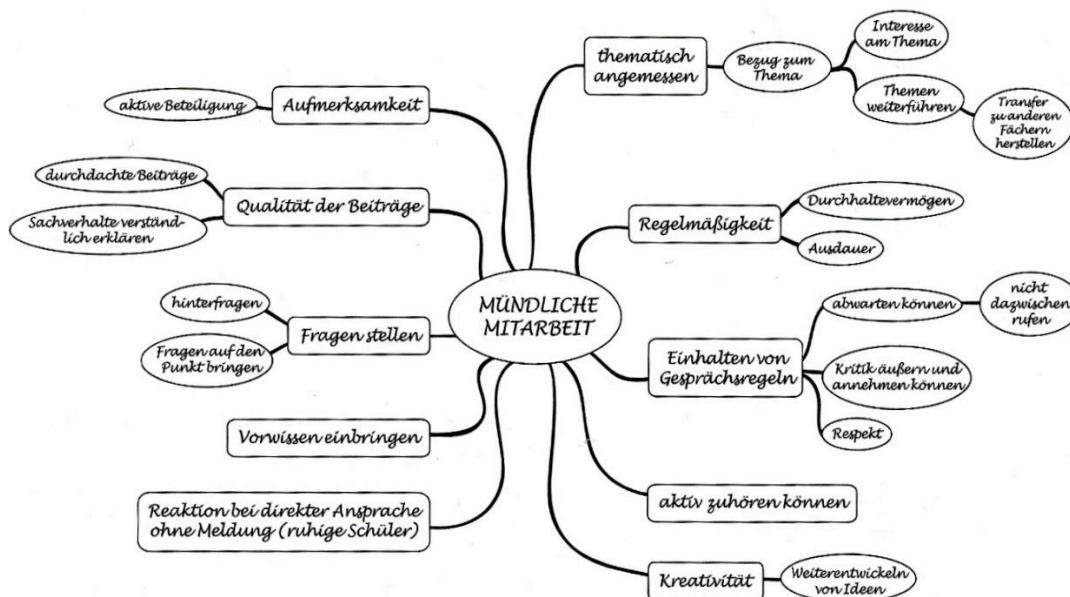


Abbildung 4: Kriterien für die Bewertung mündlicher Mitarbeit aus Krumwiede u. a. (2014, S. 24)

Ferner wird die Verwendung von Kriterienkatalogen, Beobachtungs- bzw. Selbstbeobachtungsbögen und Rückmeldebögen empfohlen². Diese sollten zuvor

¹ Siehe dazu das Kompetenzmodell nach Ulf Abraham in 3.1.5 Mündliche Leistungen.

² Eine Auswahl dieser Kataloge befindet sich im Teil A des Anhangs.

mit den Lernenden eingeübt werden. Dann können sie gute Möglichkeiten für eine differenzierte Leistungsrückmeldung bieten (vgl. ebd., S. 25).

Langhammer gibt einen Vorschlag für die Unterscheidung von drei Leistungsfeldern im Bereich des Mündlichen: Bei dem Leistungsfeld der *non-verbalen und vokalen Kommunikation* geht es vor allem um Sprechausdruck und Körpersprache.

Im zweiten Leistungsfeld der *verbalen Kommunikation* zählen die Inhalte nach Anforderungsbereichen, die Länge des Beitrags und die Grammatik.

Im dritten Leistungsfeld der *Situativen- und Gesprächs-Kommunikation* werden die Beteiligungsbereitschaft- und Häufigkeit sowie die Gesprächsfunktionen, wie z. B. Gesprächsleitung, Strukturierung oder Weiterführung, bewertet (vgl. Langhammer 1998, S. 40).

Rainer Nickel wählt eine andere Klassifizierung der Kriterien zur Bewertung mündlicher Mitarbeit, die in Tabelle 2 dargestellt wird.

Tabelle 2: Kriterien für die Bewertung mündlicher Mitarbeit nach Nickel (2004, S. 3 f.)

Kategorie	Merkmale
Lern- und Arbeitstechniken	Notizen anfertigen, verständlich (vor)lesen etc.
Kommunikationsformen	Sich zu Wort melden, aufmerksam zuhören, andere ausreden lassen, auf Mitschüler eingehen, beim Thema bleiben, Meinungen begründen etc.
Kooperation in der Lerngruppe	Regeln vereinbaren und einhalten, Konflikte zur Sprache bringen, Toleranz beweisen, sich gegenseitig helfen, höflich miteinander umgehen etc.

Zusammenfassend lässt sich konstatieren, dass es zahlreiche Varianten für die Kriterien zur Bewertung mündlicher Leistungen, insbesondere der mündlichen Mitarbeit, gibt. Dabei spielen die Begriffe der Quantität und Qualität eine große Rolle. Zudem wird besonders die Gesprächsfähigkeit, als eine der mündlichen Kompetenzen, hervorgehoben.

Letztlich bleibt es der Lehrperson, mangels rechtlicher Vorgaben, selbst überlassen, sich geeignete Kriterien, allein, zusammen mit den SuS oder im Kollegium, zu erarbeiten¹.

¹ Krieger (2003, S. 80) kritisiert, dass sich Lernende durch die mangelnden rechtlichen Vorschriften „der Willkür von Lehrern in besonderem Maße ausgesetzt“ sehen können.

3.3.2 Chancen und Grenzen

In diesem Abschnitt wird zunächst auf die Vorteile der Bewertung mündlicher Leistungen, insbesondere der mündlichen Mitarbeit, im Vergleich zu schriftlichen Leistungsüberprüfungen, wie Klassenarbeiten, eingegangen. Anschließend werden dagegen auch einige Herausforderungen bei der Bewertung mündlicher Mitarbeit beleuchtet.

3.3.2.1 Vorteile gegenüber der Bewertung schriftlicher Leistungen

Bei der Bewertung mündlicher Leistungen können sprachliche und kommunikative Leistungen berücksichtigt werden (vgl. Birkel 1984, S. 236). Während schriftliche Klassenarbeiten Leistungen eher punktuell, maximal ein bis zwei Mal pro Halbjahr, prüfen, können Leistungen im Unterrichtsgeschehen fortdauernd beobachtet werden, sodass die Bewertung der mündlichen Mitarbeit besser dazu geeignet ist, Lernprozesse kontinuierlich und in ihrer Gesamtheit zu erfassen. Durch die Möglichkeit des flexiblen Nachfragens und des Eingehens auf einzelne Lernende lässt sich die Leistung individueller feststellen als bei Klassenarbeiten, bei denen alle SuS die gleichen Aufgaben erhalten. Darüber hinaus kann eine Lehrperson im Gespräch direkte Rückmeldungen über die Richtigkeit der Beiträge bzw. den Stand im Lernprozess geben. Besonders für solche Lernende, die unter Prüfungsangst leiden und daher nicht ihre besten Leistungen in Klassenarbeiten abrufen können, bietet die Bewertung der mündlichen Mitarbeit eine Chance. Weiterhin besteht im Unterrichtsgespräch kein ständiger Zwang, die Lernenden untereinander zu vergleichen. Statt sozialer Vergleiche können SuS eher an dem individuellen Lernfortschritt gemessen werden (individuelle Bezugsnorm). Die Bewertung mündlicher Mitarbeit bietet zudem mehr Gelegenheiten für die Würdigung besonderer oder kreativer Beiträge als bei Klassenarbeiten. Das Wissen der Lernenden darüber dient außerdem als Anreiz¹ für die Beteiligung an den Unterrichtsgesprächen (vgl. Wengert 2011, S. 342 f.).

3.3.2.2 Herausforderungen

Allerdings darf die Bewertung der mündlichen Mitarbeit auch nicht als zusätzliches „Druckmittel“ benutzt werden, da die Lernenden durch Klassenarbeiten meist schon starkem Prüfungsdruck ausgesetzt werden (vgl. Nickel 2004, S. 6). Zudem haben

¹ Im Sinne der Disziplinierungsfunktion der Leistungsbewertung, siehe *Kapitel 3.1.4 Zeugnisse und Noten*.

in einer 45-minütigen Unterrichtsstunde und einer Anzahl von ca. 30 SuS nicht einmal alle Lernenden eine Chance, viel zum Unterrichtsgespräch beizutragen (vgl. Bittner 2006, S. 23 f.; Souvignier & Förster 2014, S. 7).

Die Bewertung mündlicher Leistungen stellt ein Problemfeld der schulischen Bewertungspraxis dar (vgl. Langhammer 1998, S. 38; Kirk 2004, S. 9; Krieger 2003, S. 75 ff.; Schröter 1981, S. 10; Souvignier & Förster 2014, S. 7). Die mündliche Mitarbeit und deren Bewertung werden durch zahlreiche Merkmale beeinflusst: Durch die Schülerpersönlichkeit, z. B. wie lebhaft ein Lernender auftritt, die Persönlichkeit der Lehrperson, das sich im Aufrufverhalten und der Anzahl und der Schwierigkeit von Fragen zeigt sowie durch zahlreiche situative Gegebenheiten, wie beispielsweise die Sitzordnung (vgl. Souvignier & Förster 2014, S. 11).

Schwierigkeiten bei der Beobachtung der mündlichen Mitarbeit entstehen nach Wengert (2011, S. 343 f.) zusätzlich zu den Grenzen der schulischen Beobachtungen, die in Kapitel 3.2.1.3 *Subjektive Verfahren* bereits dargestellt wurden, besonders durch die Flüchtigkeit der jeweiligen Situationen. Es gibt keine Möglichkeit, die Situation zu wiederholen oder Leistungen direkt zu vergleichen, wie das z. B. bei schriftlichen Leistungen der Fall wäre. So können, durch den zeitlichen Abstand vom Geschehen bis zur Dokumentation durch die Lehrperson, Erinnerungslücken oder sogar Verzerrungen des Wahrgenommenen entstehen. Dennoch ist die Lehrperson rechtlich nicht verpflichtet, mündliche Leistungen regelmäßig zu dokumentieren¹.

Bei der Beurteilung mündlicher Mitarbeit sind besonders drei Urteilsfehler bedeutsam: Der *Haloeffekt*, der *Rosenthal-Effekt* und der *Milde-Effekt*². Es besteht die Gefahr, dass z. B. eine hohe Quantität der Unterrichtbeiträge den Gesamteindruck über den Lernenden derart überstrahlt, dass die Beiträge auch qualitativ höher eingeschätzt werden (*Haloeffekt*). Zudem haben sich Lehrkräfte durch die schriftlichen Leistungen der SuS schnell ein Bild über den Leistungsstand der Lernenden gemacht. Problematisch wird es, wenn sie diese Erfahrungen in zukünftige (Leistungs-) Erwartungen im mündlichen Bereich übertragen. Die Folge ist, dass die Bewertung der mündlichen Mitarbeit meist in der Nähe der Bewertung der Klassenarbeiten rangiert und die mündliche Mitarbeit so nur noch als Ausschlag zum Auf- oder Abrunden der Klassenarbeitsnote fungiert (*Rosenthal-Effekt*). Hinzu kommen die möglichen Hemmungen einzelner Lehrpersonen, Noten schlechter als

¹ Siehe dazu *Kapitel 2. Rechtliche Vorgaben*.

² Diese Urteilsfehler wurden bereits erläutert in *Kapitel 3.2.2.2 Herausforderungen beim Beurteilen und Bewerten*.

„ausreichend“ zu vergeben (*Milde-Effekt*). Insgesamt ist die Gefahr subjektiver Einflüsse bei der Bewertung mündlicher Mitarbeit groß, da diese nicht „ohne Ansehen der Person“ geschehen kann (vgl. Wengert 2011, S. 343 f.). Ferner können *Reihungsfehler* genau dann entstehen, wenn eine Lehrperson im Großgruppengespräch die unmittelbar aufeinander folgenden Beiträge unterschiedlicher SuS miteinander vergleicht. Insbesondere bei der Weiterreichung einer Frage ergeben sich unterschiedliche Chancen der Beantwortung für den ersten im Vergleich zum letzten durch mehr Zeit zum Nachdenken oder zum Nachschlagen in den Unterlagen oder durch zusätzliche Hinweise der Lehrperson, (vgl. Birkel 1984, S. 234 f.).

Martenstein (2011) kritisiert im „ZEITmagazin“ die Bewertung der mündlichen Mitarbeit insofern, dass „Dauerlabern“ besonders gewürdigt und die Unterschiedlichkeit der Lernenden bezüglich Intro- bzw. Extraversion nicht berücksichtigt werde. Auch Nickel (2004, S. 7) bezeichnet das indirekte Ermutigen leistungsstarker Lernenden zur Selbstdarstellung und Demonstration von Überlegenheit als unerwünschte Nebenwirkung.

Unter sozialer Kompetenz versteht Kanning (2002, S. 155) die „Gesamtheit des Wissens, der Fähigkeiten und Fertigkeiten einer Person, welche die Qualität eigenen Sozialverhaltens – im Sinne der Definition sozial kompetenten Verhaltens – fördert“. Die vielen Faktoren, um soziale Kompetenzen zu beschreiben, können in vier Kategorien klassifiziert werden: Die *soziale Orientierung* umfasst Aspekte wie Prosozialität, die Fähigkeit zur Perspektivübernahme, Kompromissbereitschaft und das Zuhören. Die *soziale Offensivität* beschreibt die Durchsetzungsfähigkeit, Konfliktbereitschaft, Extraversion und Entscheidungsfreudigkeit einer Person. Dazu kommen die Kategorien der *Selbststeuerung* und die der *Reflexivität* (vgl. Kanning 2009, S. 194 f.).

Eine aktuelle Studie von Jurkowski und Hänze (2016) bestätigt, dass die Offensivität eines Lernenden, welche sich u. a. durch selbstbewusstes und extrovertiertes Auftreten äußert, ein Prädiktor für eine positive Bewertung der mündlichen Mitarbeit ist.

Aus dem oben genannten Vorteil andauernder Beobachtung und Bewertung mündlicher Leistungen im Unterricht, gegenüber punktueller Leistungsfeststellung durch Klassenarbeiten, entsteht nach Weinert eine weitere Herausforderung: Die Lernenden befinden sich in einer permanenten Prüfungssituation¹, in der es gilt

¹ Weinert verwendet den Begriff der Leistungssituation. Da aber Lernen auch eine Leistung ist, hat sich in der Literatur der Begriff Prüfungssituation durchgesetzt.

Erfolge zu erzielen und Misserfolge zu vermeiden. In Lernsituationen, frei von Bewertung, ist es entscheidend, neugierig zu sein, Neues zu entdecken und dadurch zu lernen. Dabei können Fehler zur Erkenntnis verhelfen und damit produktiv für den Lernprozess sein. In Prüfungssituationen hingegen wird, durch das Verbleiben in einem sicheren Terrain, Korrektheit angestrebt (Weinert 1999b, S. 33). Lernen erfordert also den Mut, etwas Neues zu wagen und Fehler zu machen. Ob Fehler als motivierend oder frustrierend erlebt werden, hängt von der Situation ab. Daher ist eine positive Fehlerkultur essentiell für guten Unterricht (vgl. Weinert 1999a, S. 102 ff.).

„Erfolgreicher Unterricht braucht [...] beides, und zwar im Bewußtsein der Schüler möglichst separiert: viele entspannte Gelegenheiten zum intensiven Lernen und anspruchsvolle Leistungssituationen“ (Weinert 1999b, S. 33).

Die Forderung nach einer Trennung von Lern- und Prüfungssituation widerspricht diametral der permanenten Beurteilung der mündlichen Mitarbeit. Im Allgemeinen bemühen sich Lehrkräfte daher darum, ihren SuS das Gefühl zu geben, nicht ständig bewertet zu werden (vgl. Birkel 1984, S. 234). Darüber hinaus sollte eine Lehrperson im Unterricht „bewertungsfreie Zonen“ einrichten, in denen sich Lernende sorgenfrei ausprobieren können (vgl. Souvignier & Förster 2014, S. 10 ff.).

Ein Konsens und verbindliche Absprachen über Ziele, Vorgehensweisen und Kriterien, auf Schulebene und oder im Kollegium, würden den Lehrkräften die Bewertung der mündlichen Mitarbeit, auch in Bezug auf die Legitimation und die Rechtfertigung gegenüber den SuS und ihren Eltern, erleichtern (vgl. ebd, S. 10; Nickel 2004, S. 4).

3.3.3 Forschungsstand

Wie sieht die Bewertung der mündlichen Mitarbeit nun in der Praxis aus und wie zufrieden damit sind die Beteiligten? Unter Zuhilfenahme der verfügbaren empirischen Ergebnisse wird der Frage in diesem Abschnitt nachgegangen. Dabei ist zu beachten, dass die bisherigen Studien sich meist der mündlichen Note als Ganzes gewidmet haben.

Schröter (1981, S. 10) bezeichnet die Bewertung mündlicher Leistung als „weißen Fleck auf der Landkarte“ der pädagogischen Forschung. Weiterhin heißt es: „Bis auf Untersuchungen über Abschlußexamina [gibt es] nur einige sporadische und relativ unverbindliche Bemerkungen in der Literatur. Im wesentlichen [sic] verläßt man sich dabei noch immer auf das Naturtalent des Lehrers“. Schröter beschreibt

weiterhin die Ergebnisse zweier Befragungen, die im Rahmen von Examensarbeiten durchgeführt wurden¹:

Die 23 befragten Lehrkräfte von H. Pahlke im Fach Sachunterricht der 3. und 4. Jahrgangsstufe gaben an, dass die mündliche Mitarbeit einen Anteil an der Gesamtnote von 15-75% und im Mittel 50% einnimmt. Fünf von ihnen gaben an, nach jeder Unterrichtsstunde Plus-Minus-Zeichen zu vermerken, sechs weitere erklärten, dass sie Notizen und Bemerkungen schriftlich festhalten. Zwei Lehrkräfte lehnten die Dokumentation grundsätzlich ab (vgl. Schröter 1981, S. 11).

Bei einer weiteren Erhebung von B. Wienert 1977 mit 178 Lehrkräften des Sachunterrichts aus der gesamten Bundesrepublik wurde folgendes festgestellt: 90% der Lehrpersonen wissen zwar um die Problematik der Bewertung, jedoch scheinen die wenig konkreten Erlasse nicht dabei zu helfen, geeignete Verfahren für die Praxis zu entwickeln. In weniger als 15% der erfassten Schulen gibt es diesbezüglich Fach- und Konferenzbeschlüsse. Die häufigsten Leistungen, die als bedeutsam für die Halbjahresnote genannt wurden, sind: Antworten auf Lehrerimpulse, Darstellung eigener Gedankengänge, Entwicklung neuer Ideen, Beiträge zum Unterrichtsgespräch und die Mitarbeit zu Beginn der Unterrichtsstunde. Mehr als die Hälfte der Lehrpersonen gab an, sich während der Unterrichtsstunde Notizen über die mündlichen Leistungen der Lernenden zu machen, dabei besonders herausstechend positive oder negative Leistungen. Etwa die Hälfte gab an, sich erst nach der Stunde Anmerkungen aufzuschreiben, mit der Begründung, dass das sichtbare Notieren der Leistungen für die SuS einen „Drohcharakter“ aufweist (vgl. ebd., S. 198 f.).

Nach einer Lehrerbefragung von Schwark u. a. (1987, S. 67 ff.) gaben nur 7% der Lehrenden an einer Grundschule an, dass es Schwierigkeiten bei der Bewertung mündlicher Leistungen gibt. Dieses Ergebnis lässt sich jedoch nur bedingt auf die Praxis an Gymnasien übertragen.

In der Fragebogen-Erhebung von Langhammer (1997, S. 80 ff.) mit 44 Teilnehmenden gaben 90% der Lehrkräfte an, dass die Bewertung mündlicher Leistungen kein Bestandteil ihrer Ausbildung war.

Kirk (2004, S. 13 ff.) führte eine Befragung von Lehrkräften aus verschiedenen Schulen und Schulstufen mit verschiedenen Fächern durch². Dabei kristallisierte

¹ Weitere Angaben zum Zeitraum etc. wurden nicht gemacht.

² Dabei gibt Kirk an, dass die Befragung nicht repräsentativ ist. Es werden keine Angaben zum Zeitpunkt der Befragung oder zur Stichprobengröße gemacht.

sich heraus, dass viele unterschiedliche Leistungen zu berücksichtigen sind und die Schwerpunkte von jeder Lehrperson individuell gesetzt werden. Zudem zeigte sich in der Befragung, dass es Unsicherheiten gab, mündliche, schriftliche und fachspezifische Leistungen klar voneinander abzugrenzen. Nach der Autorin bieten die schulrechtlichen Vorgaben diesbezüglich keine ausreichende Hilfe. Durch diese Uneinheitlichkeit entsteht Verunsicherung bei den Lernenden und deren Eltern.

Auf Nachfrage nannten die Lehrenden u. a. folgende Problembereiche bei der Bewertung mündlicher Leistungen: Einbeziehung/ Berücksichtigung aller gezeigten Leistungen, die Erfassung des aktuellen Standes und die Berücksichtigung von Leistungsveränderungen, die fehlende Objektivität und Sachbezogenheit bei der Beurteilung mündlicher Leistungen, unklare Beobachtungskriterien, interkollegiale Unterschiede, die Berücksichtigung verschiedener Voraussetzungen der Lernenden (sprachliche Kompetenz) sowie die Belastung durch die Notwendigkeit einer ständigen Beobachtung.

Die Dokumentation der Leistungen erfolgt, nach den Angaben der Lehrkräfte, in überwiegend strukturierter Form mit Notenlisten, teilweise auch Beobachtungsbögen oder pädagogischen Tagebüchern. Die SuS werden unterschiedlich oft über die aktuelle Bewertung informiert. Die Spanne reicht von regelmäßig über teilweise unregelmäßig bis zu Informationen nur auf Anfrage (vgl. Kirk 2004, S. 13 ff.).

Krieger (2003, S. 81 ff.) führte zwei Studien zur Praxis der Bewertung „des Mündlichen“, im Rahmen von Seminarprojekten von 1996 bis 1998, durch. Die erste Studie basiert auf Erfahrungsberichten von 479 Studierenden, die zweite Studie auf einer Befragung von 111 Lehrerinnen und Lehrern, sodass zwei Sichtweisen auf die Bewertungspraxis miteinander verglichen werden konnten.

Die Studierenden gaben in der ersten Studie an, dass ihre Lehrkräfte in der Vergangenheit folgende Kriterien als relevant für die mündliche Note nannten: Regelmäßige Mitarbeit (40%), Qualität von Beiträgen (23%), Hausaufgaben bearbeitet (8%) und Leistungen in mündlichen Prüfungen (4%). Jedoch spielten nach Aussagen der ehemaligen Lernenden die schriftlichen Leistungen, wie z. B. Protokolle, einen großen Anteil an der mündlichen Note. Die Kriterien wurden nach 39% der Befragten offengelegt, nach 31% nicht und 30% gaben an, dass es von Lehrperson zu Lehrperson unterschiedlich war. Des Weiteren fertigten die Lehrkräfte von 52% der Studierenden Notizen an, die für die Notenbesprechung dienten. 32% der Studierenden gaben an, dass nur wenige LuL Notizen machten und nach 13% der Studierenden wurden von den LuL gar keine Notizen angefertigt.

Bezüglich der Zufriedenheit in der Praxis nannten Studierende Schwierigkeiten bei der Notenbesprechung: Oft Diskussionen, teils heftig (66%), Diskussionen nur bei bestimmten Lehrkräften (11%), meistens Konsens (18%) und keine Diskussionen (5%). Bei der Frage nach der hohen Gewichtung der mündlichen Leistungen und den damit verbundenen Vor- und Nachteilen nannten sie die Kompensationsmöglichkeit für verpatzte Arbeiten (26%), spiegelt das langfristige Niveau besser (10%), fördert Beteiligung am Unterricht (9%), Beurteilung ist subjektiv, willkürlich (42%) und Beurteilung der mündlichen Mitarbeit ist nachteilig für stillere, ängstlichere Lernende (30%).

Bei der zweiten Studie, auf Seiten der Lehrkräfte, war die Resonanz auf die Frage nach der Bekanntheit der rechtlichen Vorgaben wie folgt: Vorschriften sind bekannt (75%), nicht bekannt (15%), keine Angaben (7%) und es gibt keine Vorschriften (3%). Auf die Frage nach welchen Vorschriften sie sich explizit orientieren, gaben sie mit 21% an, sich an bekannten Vorschriften zu orientieren, mit 8% an keinen und 71% der Befragten machten keine Angaben. Krieger vermutet diesbezüglich, dass die Vorgaben zwar bekannt, jedoch nicht die bestimmten Passagen angegeben werden können. Daraus wird geschlossen, dass diese Kenntnisse nicht in der Lehrerausbildung systematisch vermittelt, sondern eher im Kollegium tradiert werden (vgl. ebd., S. 83 f.).

Die Lehrenden trafen folgende Aussagen bezüglich der Häufigkeit ihrer Dokumentation mündlicher Leistungen: Nach jeder Stunde (30%), nach jeder Woche (20%), alle zwei Wochen (4%), am Ende der Unterrichtseinheit (10%), unregelmäßig (19%), keine Notizen/ Gesamtbild entscheidend (11%) und keine Angaben (6%).

Die Kriterien für die mündliche Note verteilten sich wie folgt: Qualität der Beiträge (55%), Quantität der Beiträge (Häufigkeit 47 %, Mitarbeit 30%), Hausaufgaben und Heftführung (39%), sprachlicher Ausdruck (16%), Referate und Protokolle (12%) und Leistungen aus z. B. Vokabeltests (10%). Diese Kriterien seien den SuS bekannt (54%), nicht bekannt (8%) oder unschlüssig (38%).

Auf die Frage nach Schwierigkeiten bei der Beurteilung des Mündlichen gaben Lehrkräfte an: Objektivität bewahren, Bevorzugung/ Benachteiligung vermeiden (41%), den stillen, zurückhaltenden Lernenden gerecht werden (16%), Konflikte bei der Notenbesprechung (11%) und keine Angaben (32%). Bei der gezielten Nachfrage, ob es Probleme bei den Notenbesprechungen gibt, nannten Lehrende: Immer (6%), selten (60%), nie bzw. es gibt keine Besprechungen (21%) und keine Angaben (13%).

Krieger konstatiert, dass es erhebliche Unterschiede in der Wahrnehmung von Lernenden und Lehrenden insofern gibt, dass Lehrpersonen die Bewertung der mündlichen Note als „harmonischen, unspektakulären Vorgang“ beschreiben, während ehemalige SuS über häufige Diskussionen und Uneinigkeiten berichteten (vgl. Krieger 2003, S. 86).

Zusammenfassend ergaben die vorliegenden Studien, dass die wenig konkreten rechtlichen Vorgaben als nur bedingt geeignet empfunden wurden, um einen Rahmen für die Entwicklung praktikabler Bewertungskriterien und -verfahren zu schaffen (vgl. Schröter 1981; Kirk 2004).

Bei der Befragung von Langhammer (1997) gaben 90% der Lehrkräfte an, dass die Bewertung mündlicher Leistungen kein Teil ihrer Lehrerausbildung war. Dies trägt nach Kirk (2004) dazu bei, dass jede Lehrperson individuelle Maßstäbe setzt, sodass eine Uneinheitlichkeit entsteht, die zu Unsicherheiten bei allen Beteiligten führt.

Bei den Kriterien werden in der Befragung von Krieger (2003) zum einen die Qualität und zum anderen die Quantität der mündlichen Beiträge genannt.

Die Dokumentation geschieht meist durch tabellarische Notenlisten. Bei der Häufigkeit der Dokumentation ist, in Bezug auf Kirk (2004) und Krieger (2003), eine große Spanne von regelmäßigen „nach jeder Stunde“ bis hin zu unregelmäßigen bis keinen Mitschriften zu bemerken. Die Studie von Krieger liefert weiterhin Erkenntnisse darüber, dass die Bewertungskriterien, nach Aussagen der Lernenden, abhängig von der Lehrperson, in unter 50% der Fällen bekannt gegeben werden. 54 % der Lehrenden gaben an, dass die Kriterien transparent dargestellt werden.

Auf Nachfrage nannten Lehrende u. a. folgende Schwierigkeiten bei der Bewertung: Die fehlende Objektivität und Sachbezogenheit, unklare Kriterien, die Berücksichtigung verschiedener Voraussetzungen der Lernenden. Lernende gaben diesbezüglich ähnliche Schwierigkeiten an: Die Beurteilung ist, aus Sicht der SuS, subjektiv, willkürlich und nachteilig für stillere SuS. Des Weiteren ist die Diskrepanz zwischen der Wahrnehmung der Lehrenden, als harmonischen Vorgang mit der Einigkeit über die Bewertung, und der Wahrnehmung der Lernenden, als konfliktbehafteter Vorgang, bemerkenswert.

Insgesamt gab es seit Schröter (1981) von dem „weißen Fleck auf der Landkarte“ der pädagogischen Forschung gesprochen hat, nur wenige weitere Studien, die sich dem Thema gewidmet haben. Daher bleibt dieser Bereich weiterhin ein relativ unerforschtes Feld.

3.4 Zusammenfassung und Ausblick

Zusammenfassend werden unter schulischen Leistungen die von der Schule initiierten Lernprozesse und -ergebnisse der SuS verstanden. Dabei geht es, neben dem Ergebnis, explizit auch um die Ausführung einer zielgerichteten Tätigkeit, was besonders für die mündliche Mitarbeit zentral ist. Die mündliche Mitarbeit ist ein wichtiger Teil mündlicher (Schul-) Leistungen, da die mündlichen Beiträge zum Unterrichtsgespräch essentiell für einen anregenden und lernförderlichen Unterricht sind. Dafür ist vor allem die Gesprächsfähigkeit, als eine der mündlichen Kompetenzen, eine elementare Voraussetzung.

Aufgrund des geltenden Leistungsprinzips bei der Verteilung von Lebenschancen, besteht die Notwendigkeit einer Klassifizierung der jeweiligen Ausprägungen der erbrachten Leistungen, die sich in den verschiedenen Abstufungen der Noten widerspiegeln. Dafür müssen schulische Leistungen beobachtet, anschließend anhand eines Referenzrahmens beurteilt und bewertet werden.

Die Bewertung mündlicher Mitarbeit, als grundlegender Bestandteil der mündlichen Note, ist ein Problemfeld der schulischen Leistungsbewertung, da sie besonders durch die Flüchtigkeit gekennzeichnet ist.

Zudem ergeben sich bei den „subjektiven Verfahren“ der Leistungsbewertung Schwierigkeiten bei der Gewährleistung der Testgütekriterien, was insbesondere bei der Bewertung mündlicher Mitarbeit eine Herausforderung darstellt. Denn schriftliche und mündliche Leistungen weisen strukturelle Unterschiede auf, wie in *Kapitel 3.1.5 Mündliche Leistungen* dargestellt wurde. Außerdem gibt es, vor dem Hintergrund der wenig konkreten rechtlichen Vorgaben, zahlreiche Herausforderungen: Bei der Festlegung geeigneter Kriterien und der Berücksichtigung zurückhaltender SuS sowie beim Widerspruch der gewollten kontinuierlichen Bewertung einerseits und der für das Lernen hinderlich wirkenden konstanten Prüfungssituation andererseits.

In den wenigen vorliegenden Studien zu dem Thema wird eine Diskrepanz zwischen Schüler- und Lehrerwahrnehmung deutlich insofern, dass LuL den Vorgang als überwiegend harmonisch bezeichneten, wohingegen ehemalige SuS von Uneinigkeiten berichteten.

Die rechtlichen Vorgaben sind den LuL zwar bekannt, können aber teilweise nicht benannt werden, was sich auch durch Schwierigkeiten bei den Begrifflichkeiten zeigte.

Bei der Häufigkeit der Dokumentation der mündlichen Mitarbeit gibt es eine weite Spanne von sehr regelmäßig bis unregelmäßig oder keine.

Die Qualität und Quantität (dabei Anzahl und Regelmäßigkeit) scheinen zentrale Kriterien in der Praxis der Bewertung mündlicher Mitarbeit zu sein.

Es werden einige Forschungslücken deutlich: Die Vielzahl der möglichen Kriterien zur Bewertung der mündlichen Mitarbeit spiegelt sich in den Studien kaum wider. Aspekte, wie die Extraversion oder die Gesprächsfähigkeit, werden, wenn überhaupt, als Schwierigkeit bei der Bewertung genannt. Eine detailliertere Befragung, gezielt zur mündlichen Mitarbeit, über weitere Kriterien (zusätzlich zur Qualität und Quantität) und deren Gewichtung ist daher erforderlich.

Bezüglich der Dokumentation bleibt die Zufriedenheit der Lehrkräfte mit der Praxis zu klären. Ungeklärt ist ferner die Sensibilität der Lehrpersonen für den hohen Anspruch an Leistungsbewertung, der u. a. in *Kapitel 3.2.1.2 Gütekriterien* deutlich wurde und wie sie ihre Bewertung vor diesem Hintergrund selbst einschätzen.

Des Weiteren befassten sich die vorliegenden Studien nicht mit den möglichen Unterschieden zwischen Lehrkräften aus verschiedenen domänenspezifischen Aufgabenfeldern¹ oder mit unterschiedlich langer Berufserfahrung.

Insgesamt handelt es sich um ein noch wenig erforschtes Themenfeld ohne vorherige Hypothesenbildung. Daher sollen erst einmal, durch einen eher explorativen Zugang, die Voraussetzungen geschaffen werden, um die Praxis der Bewertung mündlicher Mitarbeit genauer beschreiben zu können.

¹ In der gymnasialen Oberstufe wird zwischen vier Aufgabenfeldern unterschieden: 1. Sprachlich-literarisches Aufgabenfeld, 2. Gesellschaftswissenschaftliches Aufgabenfeld, 3. Mathematisch-naturwissenschaftlich-technisches Aufgabenfeld und 4. Sport (vgl. HSchG § 5, S.19).

4. Studie „Die Bewertung mündlicher Mitarbeit im Unterricht“

Ausgehend von den theoretischen Grundlagen soll diese Arbeit, mithilfe einer populationsbeschreibenden Studie, den oben genannten Forschungslücken nachgehen und so einen besseren Einblick in das Forschungsgebiet ermöglichen. Zudem wird untersucht, ob die bereits vorliegenden Ergebnisse bestätigt werden können.

4.1 Forschungsfragen

Es werden drei übergeordnete Forschungsfragen untersucht, die mehrere Teilfragen enthalten. Dazu gehören Fragen zu den Kriterien, der Dokumentation und zu den Schwierigkeiten bei der Bewertung der mündlichen Mitarbeit. Dabei werden auch potenzielle Unterschiede zwischen den Aufgabenfeldern oder bei Lehrpersonen mit unterschiedlicher Berufserfahrung herausgearbeitet.

1. Nach welchen Kriterien bewerten die LuL die mündliche Mitarbeit?
 - 1.1. Welche Kriterien werden von den Lehrkräften benannt?
 - 1.2. Wie werden vorgegebene Kriterien gewichtet?
 - 1.3. Wird die, bei den SuS unterschiedlich ausgeprägte, Extraversion berücksichtigt?
 - 1.4. Inwiefern spielt die Gesprächsfähigkeit eine Rolle?
 - 1.5. Beeinflussen schriftliche Leistungen die Bewertung mündlicher Mitarbeit?
 - 1.6. Gibt es Unterschiede zwischen den Aufgabenfeldern bei der Rangfolge der Kriterien?
 - 1.7. Gibt es Unterschiede zwischen Lehrkräften mit mehr oder weniger Berufserfahrung bei der Rangfolge der Kriterien?
 - 1.8. Sind die Lehrkräfte bei der Bewertung mündlicher Mitarbeit auf sich allein gestellt oder gibt es Vorgaben an den Schulen oder Absprachen im Kollegium?
 - 1.9. Zu welcher Bezugsnorm tendieren die Lehrkräfte? Gibt es dabei Unterschiede zwischen Lehrkräften mit mehr oder weniger Berufserfahrung?

2. Wie wird die mündliche Mitarbeit dokumentiert?
 - 2.1. Welche Methode wird verwendet?
 - 2.2. Wie häufig findet die Dokumentation statt?
 - 2.3. Gibt es Unterschiede zwischen Lehrkräften mit mehr oder weniger Berufserfahrung?
3. Wie schätzen die LuL die Schwierigkeiten bei der Bewertung mündlicher Mitarbeit ein?
 - 3.1. Werden Schwierigkeiten von den Lehrkräften benannt?
 - 3.2. Für wie objektiv, reliabel und valide halten sie ihre Bewertung? Wie schätzen sie die Transparenz der Kriterien und die Akzeptanz der Bewertung ein?

4.2 Methode

Im Folgenden wird die Untersuchungsmethode anhand der Stichprobe, des Forschungsdesigns und der Durchführung der Studie erläutert.

4.2.1 Stichprobe

Aus der Grundgesamtheit der Hessischen Lehrkräfte an Gymnasien wurde folgende Stichprobe untersucht: Fünfzig Lehrkräfte ($n_{gesamt} = 50$) aus sieben Gymnasien aus Kassel und dessen Landkreis. Die Stichprobe umfasst 60% Lehrerinnen und 40% Lehrer. In dieser Arbeit wird jedoch nicht weiter auf geschlechtsspezifische Unterschiede eingegangen. Die Auswahl fand eher willkürlich bis zufällig statt, da ausschließlich Lehrkräfte aus Kassel und Umgebung für diese Befragung hinzugezogen wurden und die Teilnahme freiwillig war. Zudem gibt es keine Einschränkung hinsichtlich der Fächer oder der Berufserfahrung der Lehrkräfte. Die befragten LuL gaben (meist je zwei) Fächer an, die sie unterrichten. Diese können aus Abbildung 5 entnommen werden:

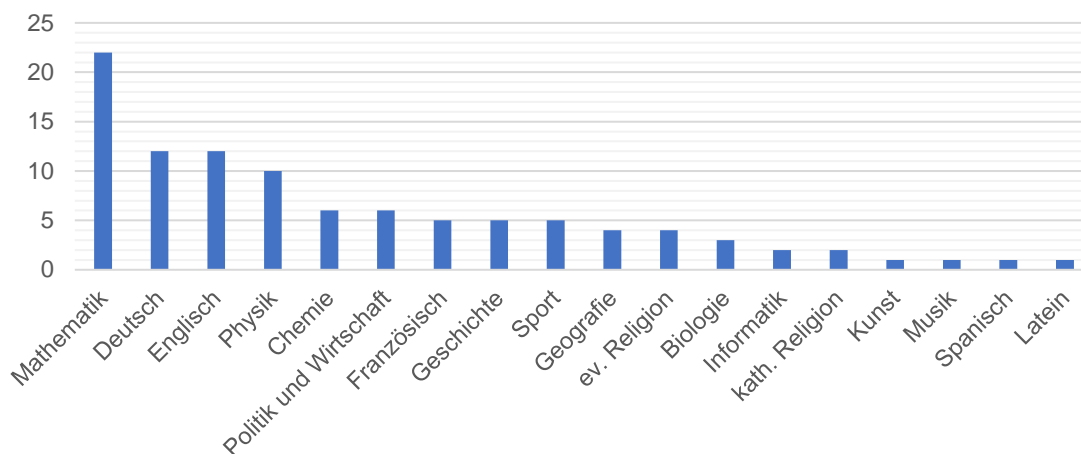


Abbildung 5: Verteilung der befragten Lehrkräfte nach Fächern

Daraus ergibt sich folgende Verteilung nach Aufgabenfeldern, die nach HSchG § 5 festgelegt ist:

Das Sprachlich-literarische Aufgabenfeld (AF1) mit den Fächern Deutsch, Englisch, Französisch, Italienisch, Spanisch, Russisch, Latein, Griechisch, Kunst und Musik, das Gesellschaftswissenschaftliche Aufgabenfeld (AF2) mit den Fächern Geschichte sowie Politik und Wirtschaft, Wirtschaftswissenschaften, Erdkunde, katholische und evangelische Religionslehre, Ethik und Philosophie, das Mathematisch-naturwissenschaftlich-technische Aufgabenfeld (AF3) mit den Fächern Mathematik, Biologie, Chemie, Physik und Informatik und das Aufgabenfeld Sport (AF4).

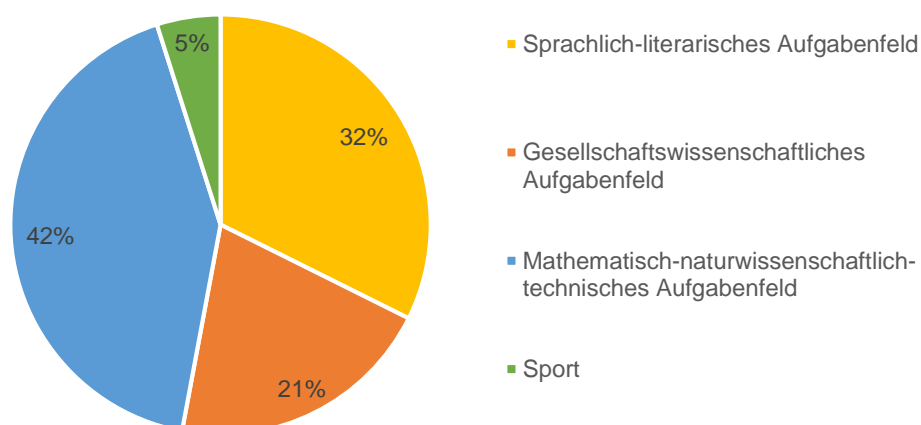


Abbildung 6: Verteilung der befragten Lehrkräfte nach Aufgabenfeldern

Die Unterscheidung nach Aufgabenfeldern, die in Abbildung 6 dargestellt ist, wird später im Abschnitt 4.3 *Ergebnisse* eine Rolle spielen.

Ebenso sei die Verteilung der Lehrkräfte nach Berufserfahrung in Abbildung 7 angeführt. Hierbei werden folgende Kategorien unterschieden: Lehrkräfte im Vorbereitungsdienst (LIV), mit 1-10 Jahren Berufserfahrung, 11-20 Jahren

Berufserfahrung, 21-30 Jahren Berufserfahrung und mehr als 30 Jahren Berufserfahrung.

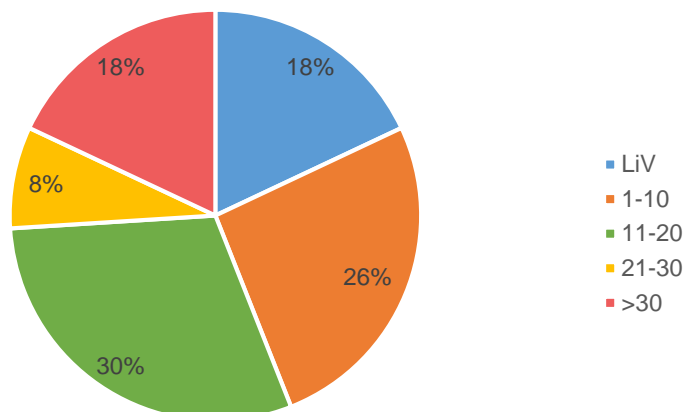


Abbildung 7: Verteilung der befragten Lehrkräfte nach Berufserfahrung in Jahren

Diese Kategorien werden zu folgenden drei Lehrergruppen klassifiziert: Lehrkräfte mit wenig Berufserfahrung (LiV, 1-10 Jahre) 44%, Lehrkräfte mit mittlerer Berufserfahrung (11-20 Jahre) 30% und Lehrkräfte mit hoher Berufserfahrung (21-30 Jahre, >30 Jahre) 26 %. Die Klassifizierung dient der Vereinfachung für die spätere Auswertung und erfolgte danach, dass sich möglichst ähnlich große Gruppen ergeben.

4.2.2 Design und Durchführung

Die teilnehmenden Lehrkräfte wurden einzeln, mithilfe einer schriftlichen und anonymen Online-Befragung¹, die im Zeitraum vom 21.06.-05.07.2016 durchgeführt wurde, befragt. Es handelte sich dabei um ein Querschnittsdesign. Die Studie wurde nicht-experimentell konstruiert und diente der Bestandsaufnahme der gegenwärtigen Bewertungspraxis.

Für eine Übersicht aller Fragen sowie die Gestaltung des Fragebogens sei auf den Teil B des Anhangs verwiesen. Der Vorteil der Online-Befragung lag vor allem darin, dass die Fragen auf mehrere verschiedene Seiten angeordnet werden konnten. Dadurch waren nachfolgende Fragen und dessen Antwortmöglichkeiten für den Teilnehmenden vorher nicht einsehbar und konnten somit die aktuelle Frage nicht beeinflussen. Zudem war es möglich, je nach Antwortverhalten des Teilnehmenden, irrelevant gewordene Fragen automatisch auszublenden. Die mit rotem Sternchen gekennzeichneten Fragen waren sogenannte Pflichtfragen, ohne die die Befragung nicht beendet werden konnte. Die übrigen waren optional zu beantwortende Fragen.

¹ Die Befragung wurde online über das Portal auf <https://www.umfrageonline.com/> durchgeführt.

Die Befragung erfolgte in überwiegend standardisierter Form mit Rating- oder Rankingskalen.

Zu Beginn wurden Fragen zur Person bezüglich Geschlecht, Schule, Berufserfahrung und zu den unterrichtenden Fächern gestellt. Anschließend sollte anhand weiterer Fragen ermittelt werden, ob und inwiefern Vorgaben der Schulen einerseits und Absprachen im Kollegium andererseits zur Bewertung mündlicher Mitarbeit vorhanden sind¹.

Anschließend wurden die Lehrenden in einer offenen Frage nach ihren verwendeten Kriterien zur Bewertung mündlicher Mitarbeit befragt. Erst danach (auf einer neuen Seite) sollten die Teilnehmenden vorgegebene Kriterien, in Bezug auf eines ihrer Fächer, nach der Relevanz in eine Rangfolge bringen. Um einer Beeinflussung durch eine bestehende Reihenfolge vor der Beantwortung der Frage entgegenzuwirken, wurde die Reihenfolge der Kriterien bei jedem Teilnehmenden zufällig generiert. Die vorgeschlagenen Kriterien umfassten, in Bezug auf *Kapitel 3.3 Mündliche Mitarbeit feststellen, beurteilen und bewerten*, die Qualität der Beiträge, Quantität (Anzahl und Kontinuität), die Gesprächsfähigkeit, die Berücksichtigung der Verschiedenheit der SuS (hier konkret die Extraversion) und die Berücksichtigung der individuellen Entwicklung des Lernenden (als grundsätzliches Kriterium). Des Weiteren wurden drei Kriterien ergänzt, die explizit kein Teil mündlicher Mitarbeit sind, deren mögliche Bedeutung für die Bewertung jedoch ebenfalls untersucht werden sollte: Die Berücksichtigung des Sozial- bzw. Störverhaltens², die Orientierung an schriftlichen Leistungen und die Erledigung der Hausaufgaben. Die Frage nach den Kriterien stellt einen Schwerpunkt dieser Studie dar.

Außerdem wurden die Teilnehmenden nach ihrer Methode und der Häufigkeit der Dokumentation, inklusive ihrer Zufriedenheit damit, befragt. Eine weitere Frage bezog sich auf die von den Lehrpersonen präferierte Bezugsnormorientierung, wobei auch Mischformen möglich waren.

Zum Schluss sollte ermittelt werden, wie die Lehrenden, vor dem Hintergrund der oben beschriebenen Herausforderungen und des hohen Anspruchs, ihre eigene Bewertung selbst einschätzen. Dies geschah, indem sie diese anhand einer Ratingskala durch eine Umschreibung der Objektivität, Reliabilität und Validität

¹ Für diese Erläuterungen diente ein freies Textfeld.

² Die Bewertung des Arbeits- und Sozialverhaltens erfolgt nach § 73 des Hessischen Schulgesetzes separat durch die Klassenkonferenz (siehe dazu *Kapitel 2*).

sowie Transparenz und Akzeptanz¹ beurteilten. Die Entwicklung eines geeigneten Messinstruments, um die tatsächliche Qualität der Bewertung zu ermitteln und diese mit den Einschätzungen der LuL zu vergleichen war in diesem Kontext natürlich nicht möglich. Zusätzlich zu der Ratingskala bot ein optionales Feld die Gelegenheit, mögliche Schwierigkeiten zu benennen.

4.3 Ergebnisse

Im Vorfeld hatten sich 73 Lehrkräfte bereit erklärt, an der Studie teilzunehmen. Davon haben 59 LuL an der Befragung teilgenommen, wobei 50 davon den Online-Fragebogen bis zum Ende bearbeitet haben. Für eine sinnvolle Auswertung wurden daher nur die Antworten dieser 50 Lehrkräfte berücksichtigt, sodass sich die Stichprobe ergab, die oben bereits beschrieben wurde.

In diesem Abschnitt werden nun die Ergebnisse, analog zur Reihenfolge der übergeordneten Forschungsfragen, dargestellt.

4.3.1 Kriterien

In diesem Abschnitt werden alle Ergebnisse bezüglich der Kriterien zur Bewertung mündlicher Mitarbeit beschrieben.

4.3.1.1 Benannte Kriterien

Die befragten LuL konnten in freien Textfeldern bis zu zehn Kriterien formulieren, die sie als relevant für die Bewertung mündlicher Mitarbeit erachten.

Um sich der Forschungsfrage 1.1 zu widmen, wurden die vielfältigen Einträge zu folgenden Hauptkategorien gruppiert: *Qualität*, *Quantität*, *Gesprächsfähigkeit*, *Weitere* und *Kein Teil der mündlichen Mitarbeit*. Die Verteilung der einzelnen Nennungen ist in Abbildung 8 dargestellt. Die Zahlen in der Legende sowie im Ringdiagramm zeigen die Anzahl der Nennungen des jeweiligen Merkmals.

¹ Unter Akzeptanz wird hier die Reibungslosigkeit des Ablaufes bzw. die Übereinstimmung mit den Lernenden verstanden.

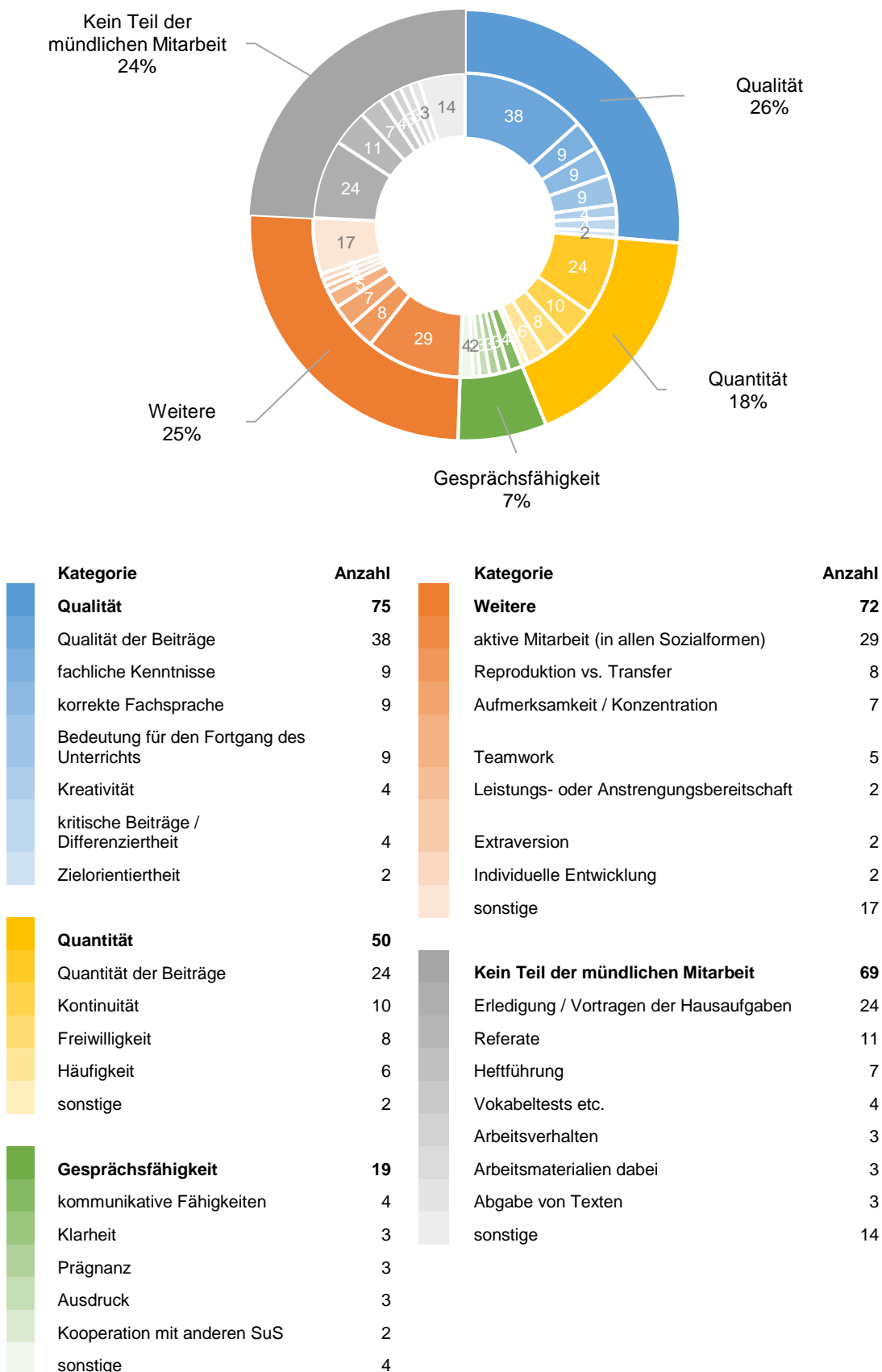


Abbildung 8: Benannte Kriterien für die Bewertung mündlicher Mitarbeit

Der mit 26 % größte Anteil der Nennungen machte das Merkmal der *Qualität* aus, bei der mehrheitlich der Begriff „Qualität“ selbst, aber auch fachliche Kenntnisse

und Fachsprache, die Bedeutung des Wortbeitrags für den Fortgang des Unterrichts sowie die Kreativität oder Differenziertheit genannt wurden.

Außerdem wurden zu 18% die Bedeutung der Häufigkeit, Kontinuität oder Freiwilligkeit der Beiträge genannt, die unter der Kategorie *Quantität* zusammengefasst werden.

7% der genannten Kriterien umfassten Merkmale, die im weitesten Sinne der *Gesprächsfähigkeit* zugeordnet werden können. Dabei wurden kommunikative Fähigkeiten, Klarheit, Prägnanz und Ausdruck bei Wortbeiträgen sowie kooperative Fähigkeiten angegeben.

In der Kategorie *Weitere* mit 25% wurden einige Merkmale zusammengefasst, wie die aktive Mitarbeit in allen Sozialformen, die Unterscheidung wiederholender, reproduktiver Beiträge von Transferleistungen, die Berücksichtigung von Aufmerksamkeit bzw. Konzentration, die Fähigkeit im Team zu arbeiten, die Leistungsbereitschaft, die Extraversion und die individuelle Entwicklung des Lernenden.

24% der genannten Kriterien umfassten Merkmale, wie die Erledigung bzw. das Vortragen der Hausaufgaben, die Qualität von Referaten, der Heftführung und von abgegebenen Texten, das Abschneiden in Vokabeltests, das Mitbringen der nötigen Arbeitsmaterialien sowie das Arbeitsverhalten. Diese können zwar Teil der mündlichen Note, jedoch nicht der mündlichen Mitarbeit, sein, wie in *Kapitel 3.1.5 Mündliche Leistungen* gezeigt wurde.

4.3.1.2 Stellenwert der Kriterien

Die Frage nach dem jeweiligen Stellenwert einzelner vorgegebener Kriterien (Forschungsfrage 1.2) für die Bewertung mündlicher Mitarbeit wurde mit einer Rankingskala untersucht. Mit der von den Lehrkräften erstellten Rangfolgen (je Fach) haben sie Punkte vergeben. Die Kodierung geschah folgendermaßen: Die insgesamt neun Kriterien wurden auf Rangplätze 1-9 verteilt, die mit 8-0 Punkten ausgewertet wurden.

Damit ergab sich für die gesamten Antworten ($n_{\text{Antworten gesamt}} = 102$) die folgende Rangfolge (mit Mittelwert und Standardabweichung):

Qualität der Beiträge (6.50; 2.49), Kontinuität der Beiträge (5.67; 1.80), Anzahl der Beiträge (4.66; 1.90), Gesprächsfähigkeit (4.20; 2.11), (Lern-) Entwicklung (3.74; 2.23), Hausaufgaben (3.40; 1.96), Gesamteindruck aufgrund schriftlicher Leistungen (2.84; 2.55), Extraversion (2.53; 2.29) sowie Sozial- und Störverhalten (2.47; 2.32).

Diese Rangfolge wird in Abbildung 9 veranschaulicht. Sie zeigt die Mittelwerte der einzelnen Kriterien und die Standardabweichung.

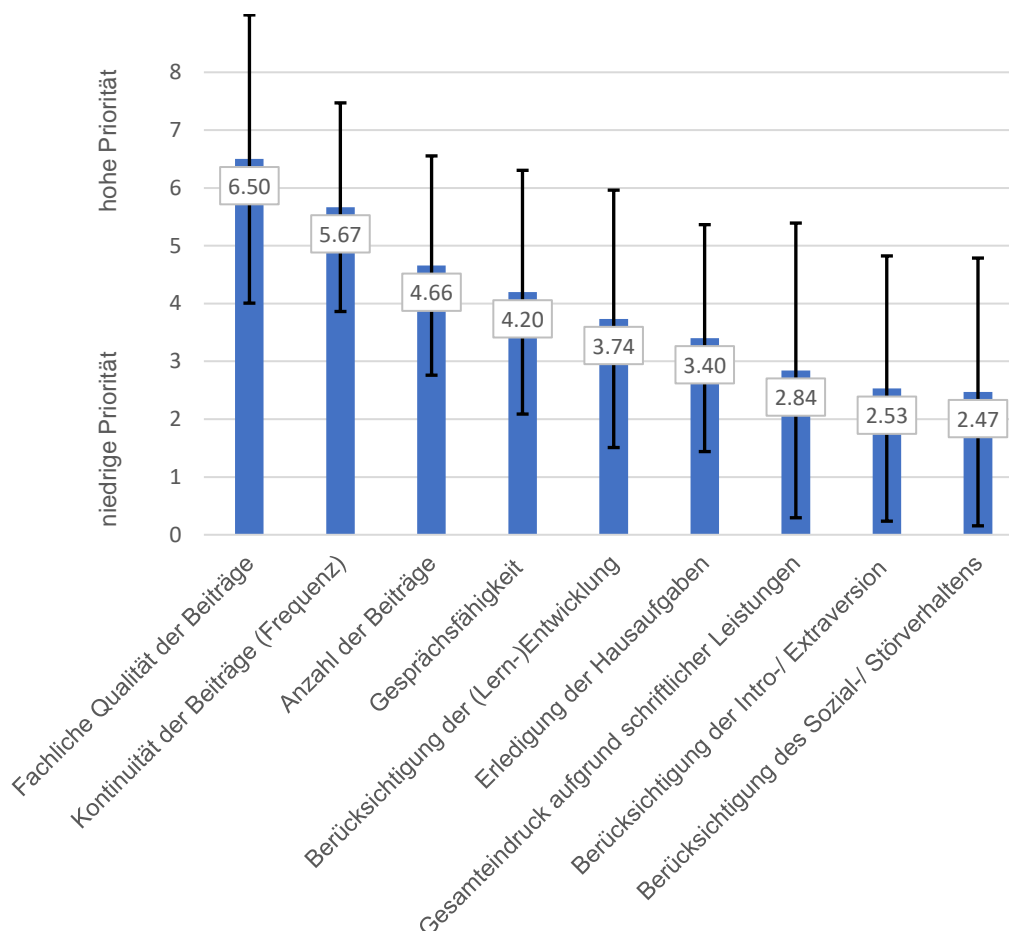


Abbildung 9: Rangfolge der Kriterien der gesamten Antworten der Stichprobe

Die individuell ausgeprägte *Extraversion* wurde bei der Frage nach den Kriterien lediglich von zwei Lehrkräften benannt. In diesem Ranking gehört die Berücksichtigung der *Extraversion* zu den am wenigsten wichtig eingestuften Kriterien (Forschungsfrage 1.3).

Die *Gesprächsfähigkeit* (Forschungsfrage 1.4) wird zwar von den Lehrkräften nicht direkt benannt und zugehörige Aspekte mit 7% nur vereinzelt angeführt, aber in diesem Ranking mit vorgegebenen Kriterien messen die Lehrkräfte der *Gesprächsfähigkeit* eine relativ große Bedeutung zu. Sie wird direkt nach Kriterien der Qualität und Quantität, bestehend aus Kontinuität und Anzahl, eingeschätzt.

Die Beeinflussung durch vorherige schriftliche Leistungen wird von den Lehrkräften selbst nicht als Kriterium für die Bewertung mündlicher Mitarbeit benannt. Bei der Rangfolge liegt dieses Kriterium auf dem siebten Rang (Forschungsfrage 1.5). Die Streuung ist bei diesem Merkmal am größten.

4.3.1.3 Stellenwert der Kriterien nach Aufgabenfeldern

Für die obige Rangfolge wurden Mittelwerte der gesamten Stichprobe gebildet. Nun werden die Unterschiede der oben beschriebenen Aufgabenfelder ausgewertet (Forschungsfrage 1.6). Da die Lehrenden vor der Beantwortung der Frage ein spezielles Fach angegeben haben, konnte so die Rangfolge in Abhängigkeit der Aufgabenfelder bestimmt werden. Dabei wurde die Reihenfolge der Kriterien analog zu Abbildung 9 beibehalten. Auf die Darstellung der Standardabweichung wurde, zugunsten der Übersichtlichkeit, verzichtet.

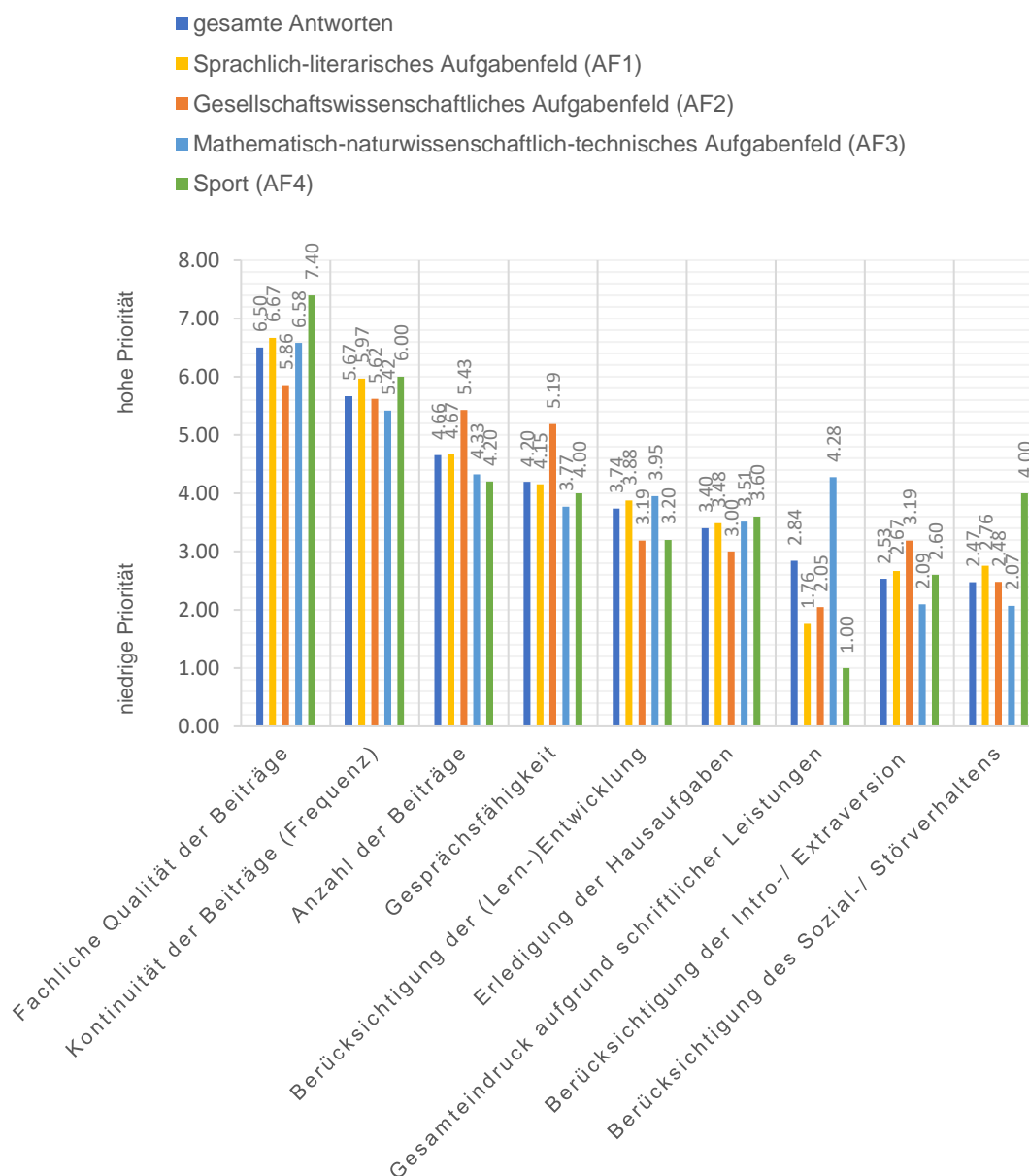


Abbildung 10: Rangfolge der Kriterien nach Aufgabenfeldern

In Abbildung 10 lassen sich durchaus Unterschiede zwischen den Aufgabenfeldern bei einigen Kriterien feststellen. Wie groß der statistische Effekt im Einzelnen ist, kann nach Cohen (1988, S. 20 ff.) durch die Effektgröße d bestimmt werden. Sie ist

durch $d = \frac{\mu_1 - \mu_2}{\sigma}$ definiert und gibt den Unterschied zwischen den Mittelwerten μ_1 und μ_2 bei gleicher Gruppengröße n und gleicher Standardabweichung σ an. Nach Cohen kann bei $d \in [0,0.2)$ von keinem Effekt, bei $d \in [0.2,0.5)$ von einem kleinen Effekt, bei $d \in [0.5,0.8)$ von einem mittleren und bei $d \in [0.8,\infty)$ von einem großen Effekt gesprochen werden.

Da hier weder die Gruppengrößen (gesamte Antworten und Antworten aus den Aufgabenfeldern) noch die Standardabweichungen übereinstimmen, ist es nötig, eine sogenannte gepoolte Standardabweichung zu bestimmen, um die Effektstärke zu ermitteln (vgl. Ellis 2010, S. 10 ff.).

Im Folgenden werden in Bezug zu den Kriterien die Mittelwerte der einzelnen Aufgabenfelder jeweils mit dem Mittelwert der gesamten Antworten ($n_{\text{Antworten gesamt}} = 102$) verglichen, um auf diese Weise die Effekte der Abweichungen zu bestimmen. Für die Berechnungen wurden stets die exakten Werte verwendet. In den Tabellen werden die Mittelwerte und Standardabweichungen auf zwei und die Effektgröße d auf drei Nachkommastellen gerundet dargestellt. Die für diese Arbeit relevanten Unterschiede sind hervorgehoben.

Tabelle 3: Abweichung der Relevanz der Qualität der Beiträge nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	6.67	2.40	0.067
Antworten gesamt	102	6.50	2.49	
Aufgabenfeld 2	21	5.86	3.02	0.248
Antworten gesamt	102	6.50	2.49	
Aufgabenfeld 3	43	6.58	2.74	0.032
Antworten gesamt	102	6.50	2.49	
Aufgabenfeld 4	5	7.40	0.89	0.367
Antworten gesamt	102	6.50	2.49	

Wie aus Tabelle 3 entnommen werden kann, weichen die Mittelwerte bezüglich der *Qualität der Beiträge* der Aufgabenfelder 1 und 3 nicht von dem Mittelwert der gesamten Antworten ab (kein Effekt). Der Mittelwert des Aufgabenfeldes 4 weicht nach oben und der des Aufgabenfeldes 2 nach unten mit jeweils einem kleinen Effekt ab.

Tabelle 4: Abweichung der Relevanz der Kontinuität der Beiträge nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	5.97	1.67	0.171
Antworten gesamt	102	5.67	1.80	
Aufgabenfeld 2	21	5.62	1.86	0.026
Antworten gesamt	102	5.67	1.80	
Aufgabenfeld 3	43	5.42	2.07	0.131
Antworten gesamt	102	5.67	1.80	
Aufgabenfeld 4	5	6.00	2.35	0.182
Antworten gesamt	102	5.67	1.80	

Bezüglich der möglichen Abweichungen bei der Relevanz der *Kontinuität der Beiträge* im Vergleich zu allen Antworten der gesamten Stichprobe wurden bei allen Aufgabenfelder sehr niedrige Werte für d bestimmt (siehe Tabelle 4), sodass von keinen Effekten ausgegangen wird.

Tabelle 5: Abweichung der Relevanz der Anzahl der Beiträge nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	4.66	1.88	0.005
Antworten gesamt	102	4.66	1.90	
Aufgabenfeld 2	21	5.43	1.91	0.406
Antworten gesamt	102	4.66	1.90	
Aufgabenfeld 3	43	4.33	1.73	0.179
Antworten gesamt	102	4.66	1.90	
Aufgabenfeld 4	5	4.20	2.58	0.237
Antworten gesamt	102	4.66	1.90	

Die Mittelwerte des Kriteriums der *Anzahl der Beiträge* der Aufgabenfelder 1 und 3 weichen nicht von dem Mittelwert der gesamten Antworten ab (kein Effekt). Der Mittelwert des Aufgabenfeldes 2 weicht nach oben und der des Aufgabenfeldes 4 nach unten mit jeweils einem kleinen Effekt ab, wie in Tabelle 5 dargestellt ist.

Tabelle 6: Abweichung der Relevanz der Gesprächsfähigkeit nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	4.15	2.43	0.020
Antworten gesamt	102	4.20	2.11	
Aufgabenfeld 2	21	5.19	1.75	0.484
Antworten gesamt	102	4.20	2.11	
Aufgabenfeld 3	43	3.77	2.02	0.206
Antworten gesamt	102	4.20	2.11	
Aufgabenfeld 4	5	4.00	1.22	0.094
Antworten gesamt	102	4.20	2.11	

Wie in Tabelle 6 zu sehen ist, weichen die Mittelwerte bezüglich der *Gesprächsfähigkeit* der Aufgabenfelder 1 und 4 nicht von dem Mittelwert der gesamten Antworten ab (kein Effekt). Der Mittelwert des Aufgabenfeldes 3 weicht nach unten mit einem kleinen Effekt und der des Aufgabenfeldes 2 nach oben mit einem kleinen bis mittleren Effekt ab.

Tabelle 7: Abweichung der Relevanz der (Lern-) Entwicklung nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	3.88	2.03	0.066
Antworten gesamt	102	3.74	2.23	
Aufgabenfeld 2	21	3.19	2.36	0.242
Antworten gesamt	102	3.74	2.23	
Aufgabenfeld 3	43	3.95	2.47	0.095
Antworten gesamt	102	3.74	2.23	
Aufgabenfeld 4	5	3.2	2.17	0.241
Antworten gesamt	102	3.74	2.23	

Hinsichtlich der *Berücksichtigung der (Lern-) Entwicklung* ergab die Berechnung der Mittelwerte bei Aufgabenfeld 1 und 3 keinen Unterschied zum Mittelwert der gesamten Antworten, da kein Effekt festgestellt wird (siehe Tabelle 7).

Die Mittelwerte der Aufgabenfelder 2 und 4 weichen nach oben von den gesamten Antworten, mit einem jeweils kleinen Effekt, ab.

Tabelle 8: Abweichung der Relevanz der Erledigung der Hausaufgaben nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	3.48	1.92	0.042
Antworten gesamt	102	3.40	1.96	
Aufgabenfeld 2	21	3.00	1.95	0.205
Antworten gesamt	102	3.40	1.96	
Aufgabenfeld 3	43	3.51	1.87	0.057
Antworten gesamt	102	3.40	1.96	
Aufgabenfeld 4	5	3.60	3.21	0.098
Antworten gesamt	102	3.40	1.96	

Die *Erledigung der Hausaufgaben* ist bei den Aufgabenfeldern 1, 3 und 4 ähnlich bedeutsam wie bei den gesamten Antworten, lediglich bei Aufgabenfeld 2 ist sie, mit einem kleinen Effekt, weniger relevant als bei den gesamten Antworten (siehe Tabelle 8).

Tabelle 9: Abweichung der Relevanz des Gesamteindrucks aufgrund schriftlicher Leistungen nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	1.76	2.17	0.441
Antworten gesamt	102	2.84	2.55	
Aufgabenfeld 2	21	2.05	2.48	0.314
Antworten gesamt	102	2.84	2.55	
Aufgabenfeld 3	43	4.28	2.29	0.580
Antworten gesamt	102	2.84	2.55	
Aufgabenfeld 4	5	1.00	1.73	0.729
Antworten gesamt	102	2.84	2.55	

Der *Gesamteindruck*, der auch durch zuvor erbrachte *schriftliche Leistungen* entsteht, ist für Aufgabenfeld 3 mit einer mittleren Effektstärke ein wichtigeres

Kriterium als bei den anderen Aufgabenfeldern. Dieses Kriterium ist weiterhin bei den Aufgabenfeldern 1 und 2 mit einem kleinen Effekt und beim Aufgabenfeld 4 mit einem mittleren bis großen Effekt weniger bedeutsam (siehe Tabelle 9).

Tabelle 10: Abweichung der Relevanz der Intro-/ Extraversion nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	2.67	2.04	0.061
Antworten gesamt	102	2.53	2.29	
Aufgabenfeld 2	21	3.19	1.97	0.295
Antworten gesamt	102	2.53	2.29	
Aufgabenfeld 3	43	2.09	2.50	0.185
Antworten gesamt	102	2.53	2.29	
Aufgabenfeld 4	5	2.60	2.70	0.031
Antworten gesamt	102	2.53	2.29	

Bezüglich der *Berücksichtigung der Intro-/ Extraversion* (siehe Tabelle 10) hebt sich lediglich das Aufgabenfeld 2 von der Gesamtzahl der Antworten, mit einem kleinen Effekt, nach oben ab.

Tabelle 11: Abweichung der Relevanz des Sozial-/ Störverhaltens nach Aufgabenfeldern

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effektgröße d
Aufgabenfeld 1	33	2.76	2.55	0.121
Antworten gesamt	102	2.47	2.32	
Aufgabenfeld 2	21	2.48	2.34	0.002
Antworten gesamt	102	2.47	2.32	
Aufgabenfeld 3	43	2.07	2.12	0.177
Antworten gesamt	102	2.47	2.32	
Aufgabenfeld 4	5	4.00	1.00	0.671
Antworten gesamt	102	2.47	2.32	

Im Sportunterricht (Aufgabenfeld 4) wird das *Sozial-/ Störverhalten* mit einem mittleren Effekt mehr berücksichtigt als bei dem Durchschnitt aller Aufgabenfelder.

Die Aufgabenfelder 1-3 weichen nur minimal ab, sodass kein Effekt festgestellt werden kann (siehe Tabelle 11).

Zusammenfassend können, bei dem Vergleich der Antworten aus den verschiedenen Aufgabenfeldern, bei einigen Kriterien Übereinstimmungen festgestellt werden. Bei anderen Kriterien können durchaus Unterschiede zwischen den Unterrichtsfächern bei dem Stellenwert der Kriterien konstatiert werden. Zunächst werden die Gemeinsamkeiten dargestellt:

Die *Qualität der Beiträge* ist für alle Aufgabenfelder das wichtigste Kriterium zur Bewertung mündlicher Mitarbeit (jeweils Rangplatz 1). Dabei gab es meist keine Effekte bezüglich Abweichungen bis auf einen kleinen Effekt beim Fach Sport, mit einem etwas höheren Mittelwert als bei den gesamten Antworten. Auch bei der *Kontinuität der Beiträge*, *Berücksichtigung der (Lern-) Entwicklung* und der *Erledigung der Hausaufgaben* gab es annähernd übereinstimmende Mittelwerte bei den jeweiligen Aufgabenfeldern, sodass stets keine oder kleine Effekte erkennbar waren. Die Unterschiede in der sozialen Offensivität der SuS wurden außerdem bei fast allen Aufgabenfeldern übereinstimmend als wenig relevant für die Bewertung mündlicher Mitarbeit eingeschätzt (im Mittel Rangplatz 8). Lediglich bei Fächern des Aufgabenfeldes 2 (wie z. B. Geschichte, Politik und Wirtschaft etc.) wird die *Intro-/ Extraversion* mit einem kleinen Effekt höher gewichtet als im Durchschnitt (siehe Tabelle 10).

Die Unterschiede zwischen den Aufgabenfeldern stellen sich wie folgt dar: Bei den Fächern aus dem Aufgabenfeld 2 deuten die Ergebnisse daraufhin, dass die *Anzahl der Beiträge* und die *Gesprächsfähigkeit* einen größeren Einfluss auf die Bewertung haben als im Durchschnitt aller Aufgabenfelder. Dies zeigen kleine bis mittlere Effekte in Tabelle 5 und 6. Dagegen zeigt sich bei Fächern aus dem Aufgabenfeld 3 (wie z. B. Mathematik, Physik etc.) eine (kleine) gegenteilige Tendenz bezüglich der *Gesprächsfähigkeit*. Dafür sind *schriftliche Leistungen* für die Bewertung mündlicher Mitarbeit in diesem Aufgabenfeld, mit einem mittleren Effekt, bedeutsamer als bei den anderen Fächern. In den Aufgabenfeldern 1 und 2 zeigen sich diesbezüglich gegensätzliche Abweichungen mit einem kleinen Effekt. Auch im Fach Sport (Aufgabenfeld 4) sind *schriftliche Leistungen* das, mit einem Mittelwert von 1.00, am wenigsten relevante Kriterium für die Bewertung mündlicher Mitarbeit. Im Vergleich mit der gesamten Anzahl an Antworten ergab dies einen mittleren bis großen Effekt (siehe Tabelle 9). Dem im Mittel am unwichtigsten (Rangplatz 9) eingeschätzten Kriterium der *Berücksichtigung des Sozial-/ Störverhaltens* kommt im Sportunterricht stattdessen eine überdurchschnittliche Bedeutung zu, das sowohl aus Abbildung 10 als auch aus Tabelle 11 deutlich hervorgeht.

4.3.1.4 Stellenwert der Kriterien nach Berufserfahrung

Weiterhin wurde untersucht, ob es Unterschiede in der Rangfolge der Kriterien bei Lehrkräften mit mehr oder weniger Berufserfahrung gibt (Forschungsfrage 1.7). Dazu wurden die Ergebnisse der Frage nach der Rangfolge der Kriterien in Abhängigkeit von der Berufserfahrung des Teilnehmenden ausgewertet. Ein erster Rangplatz wurde mit 8, ein letzter mit 0 Punkten codiert:

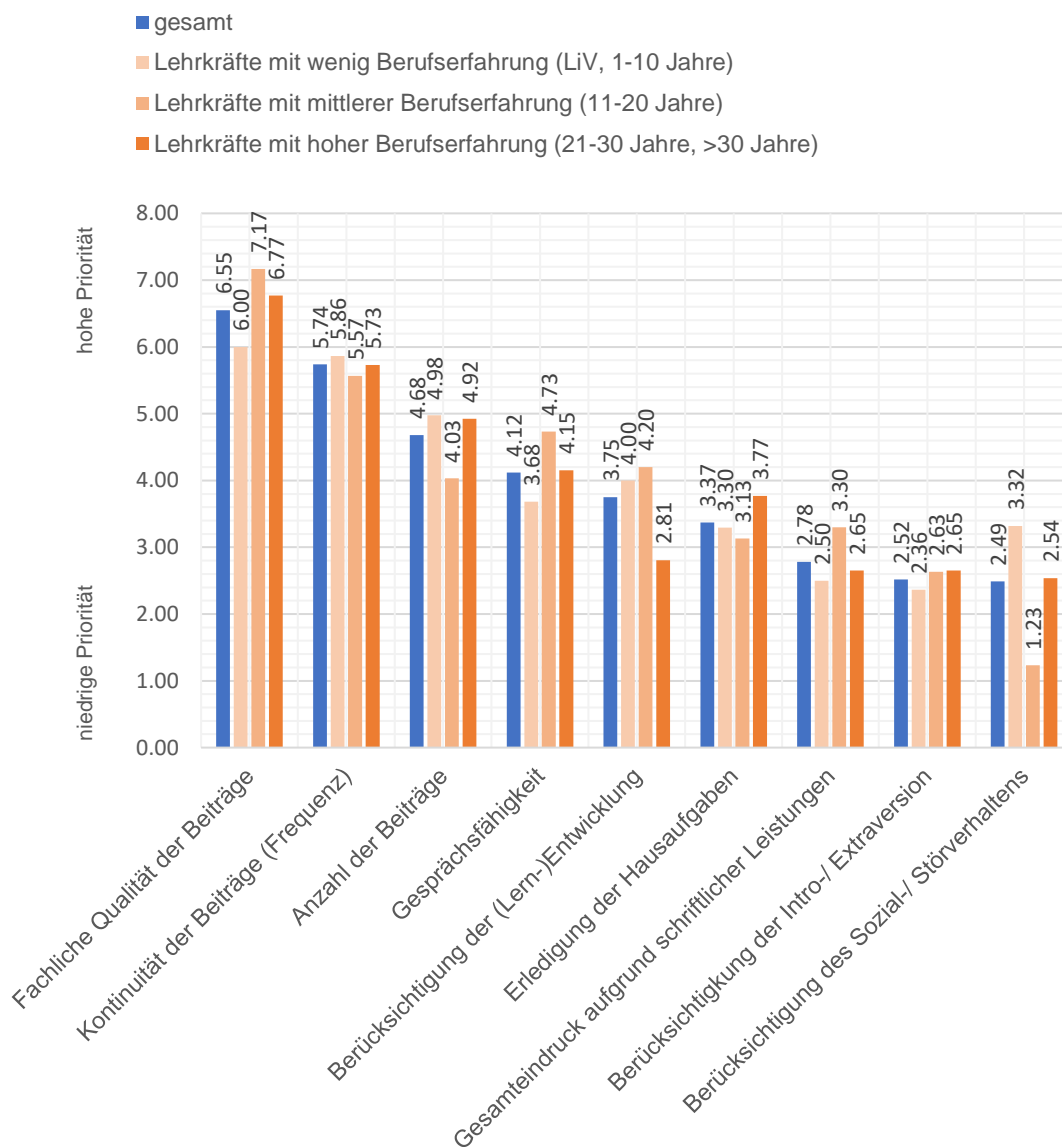


Abbildung 11: Rangfolge der Kriterien nach Berufserfahrung

Hier soll ebenfalls die Effektgröße d dabei helfen, die Abweichungen der Mittelwerte der einzelnen Gruppen von der gesamten Stichprobe einzuschätzen.

Tabelle 12: Abweichung der Relevanz der Kontinuität der Beiträge nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	5.86	2.21	0.065
Gesamte Stichprobe	50	5.74	1.73	
Mittlere Berufserfahrung	15	5.57	1.27	0.106
Gesamte Stichprobe	50	5.74	1.73	
Hohe Berufserfahrung	13	5.73	1.33	0.006
Gesamte Stichprobe	50	5.74	1.73	

Tabelle 13: Abweichung der Relevanz des Gesamteindrucks aufgrund schriftlicher Leistungen nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	2.50	2.13	0.116
Gesamte Stichprobe	50	2.78	2.54	
Mittlere Berufserfahrung	15	3.30	3.00	0.196
Gesamte Stichprobe	50	2.78	2.54	
Hohe Berufserfahrung	13	2.65	2.70	0.049
Gesamte Stichprobe	50	2.78	2.54	

Tabelle 14: Abweichung der Relevanz der Intro-/ Extraversion nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	2.36	2.30	0.068
Gesamte Stichprobe	50	2.52	2.29	
Mittlere Berufserfahrung	15	2.63	2.30	0.050
Gesamte Stichprobe	50	2.52	2.29	
Hohe Berufserfahrung	13	2.65	2.41	0.058
Gesamte Stichprobe	50	2.52	2.29	

Wie aus Tabelle 12, 13 und 14 entnommen werden kann, gibt es bei den Kriterien der *Kontinuität der Beiträge*, des *Gesamteindrucks aufgrund schriftlicher Leistungen* und der *Berücksichtigung der Intro-/ Extraversion* keine effektiven Abweichungen einzelner Gruppen mit unterschiedlicher Berufserfahrung, verglichen mit der gesamten Stichprobe ($n_{gesamt} = 50$).

Tabelle 15: Abweichung der Relevanz der Qualität der Beiträge nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	6.00	2.99	0.203
Gesamte Stichprobe	50	6.55	2.59	
Mittlere Berufserfahrung	15	7.17	2.07	0.249
Gesamte Stichprobe	50	6.55	2.59	
Hohe Berufserfahrung	13	6.77	2.36	0.086
Gesamte Stichprobe	50	6.55	2.59	

Bezüglich der *Qualität der Beiträge* (siehe Tabelle 15) sind Abweichungen mit kleinen Effekten festzustellen. Lehrkräfte mit wenig Berufserfahrung gewichten dieses Kriterien niedriger und Lehrkräfte mit mittlerer Berufserfahrung höher als im Mittel der gesamten Stichprobe.

Tabelle 16: Abweichung der Relevanz der Gesprächsfähigkeit nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	3.68	2.03	0.211
Gesamte Stichprobe	50	4.12	2.10	
Mittlere Berufserfahrung	15	4.73	2.41	0.282
Gesamte Stichprobe	50	4.12	2.10	
Hohe Berufserfahrung	13	4.15	1.78	0.017
Gesamte Stichprobe	50	4.12	2.10	

Tabelle 17: Abweichung der Relevanz der Berücksichtigung der (Lern-) Entwicklung nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	4.00	2.60	0.106
Gesamte Stichprobe	50	3.75	2.24	
Mittlere Berufserfahrung	15	4.20	1.66	0.212
Gesamte Stichprobe	50	3.75	2.24	
Hohe Berufserfahrung	13	2.81	2.03	0.429
Gesamte Stichprobe	50	3.75	2.24	

Dasselbe gilt für LuL mit mittlerer Berufserfahrung bezüglich der *Gesprächsfähigkeit* (Tabelle 16) und der *Berücksichtigung der (Lern-) Entwicklung* (Tabelle 17).

Tabelle 18: Abweichung der Relevanz der Anzahl der Beiträge nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	4.98	1.75	0.165
Gesamte Stichprobe	50	4.68	1.82	
Mittlere Berufserfahrung	15	4.03	1.54	0.367
Gesamte Stichprobe	50	4.68	1.82	
Hohe Berufserfahrung	13	4.92	2.17	0.128
Gesamte Stichprobe	50	4.68	1.82	

Tabelle 19: Abweichung der Relevanz des Sozial-/ Störverhaltens nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	3.32	2.63	0.343
Gesamte Stichprobe	50	2.49	2.31	
Mittlere Berufserfahrung	15	1.23	1.24	0.592
Gesamte Stichprobe	50	2.49	2.31	
Hohe Berufserfahrung	13	2.54	2.18	0.021
Gesamte Stichprobe	50	2.49	2.31	

Die *Anzahl der Beiträge* (Tabelle 18) ist für Lehrkräfte mit 11-20 Jahren Berufserfahrung dagegen, mit einem kleinen Effekt, weniger bedeutsam. Ebenso wird das *Sozial- /Störverhalten* als weniger wichtig für die Bewertung mündlicher Mitarbeit erachtet, mit einem mittleren Effekt, wie in Tabelle 19 dargestellt ist. Für Lehrkräfte mit wenig Berufserfahrung hingegen kommt dem *Sozial- / Störverhalten*, mit einem kleinen Effekt, bei der Bewertung eine größere Bedeutung zu.

Tabelle 20: Abweichung der Relevanz der Erledigung der Hausaufgaben nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	3.30	1.78	0.039
Gesamte Stichprobe	50	3.37	1.94	
Mittlere Berufserfahrung	15	3.13	2.00	0.121
Gesamte Stichprobe	50	3.37	1.94	
Hohe Berufserfahrung	13	3.77	2.21	0.200
Gesamte Stichprobe	50	3.37	1.94	

Lehrkräfte mit hoher Berufserfahrung berücksichtigen die (*Lern-*) *Entwicklung der SuS* (siehe Tabelle 17), mit einem kleinen Effekt, weniger stark als die gesamte Stichprobe. Dafür spielt die *Erledigung der Hausaufgaben* (siehe Tabelle 20), mit demselben Effekt, eine etwas größere Rolle für diese Gruppe.

4.3.1.5 Schulische Vorgaben

Ob es an den Schulen bestimmte Vorgaben und oder im Kollegium Absprachen zur Bewertung mündlicher Mitarbeit gibt, wurde durch eine dichotome Ja-Nein-Frage ermittelt (Forschungsfrage 1.8). Für solche Lehrenden, die mit „Ja“ geantwortet haben, bestand anschließend die Möglichkeit anzugeben, welche Vorgaben bzw. Absprachen es gibt.

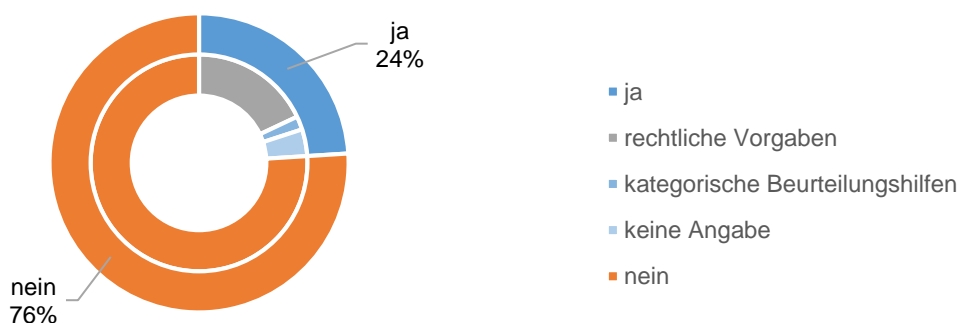


Abbildung 12: Anteil der Lehrkräfte mit Vorgaben an der Schule

Aus Abbildung 12 kann entnommen werden, dass 76% der Lehrenden angaben, dass es keine Vorgaben seitens der Schulleitung gibt. Ein großer Teil der Lehrkräfte, die diese Frage mit „Ja“ beantworten, gab auf Nachfrage rechtliche Vorgaben¹ an, die allerdings nicht unter schulische Vorgaben fallen. Eine Lehrperson beschrieb, dass sie Beurteilungshilfen erhalten hat, die aber nicht als verpflichtend gelten.

4.3.1.6 Absprachen im Kollegium

Ähnlich zeigt sich in Abbildung 13 der mit 74% große Anteil der Lehrkräfte an der gesamten Strichprobe, bei dem die Lehrenden untereinander keine Absprachen über die Bewertung mündlicher Mitarbeit treffen. 8% der Lehrkräfte berichteten von gemeinsam entwickelten Kriterienkatalogen, die unverbindlich genutzt werden können oder von solchen Absprachen im Kollegium bezüglich der Gewichtung einzelner Leistungen.

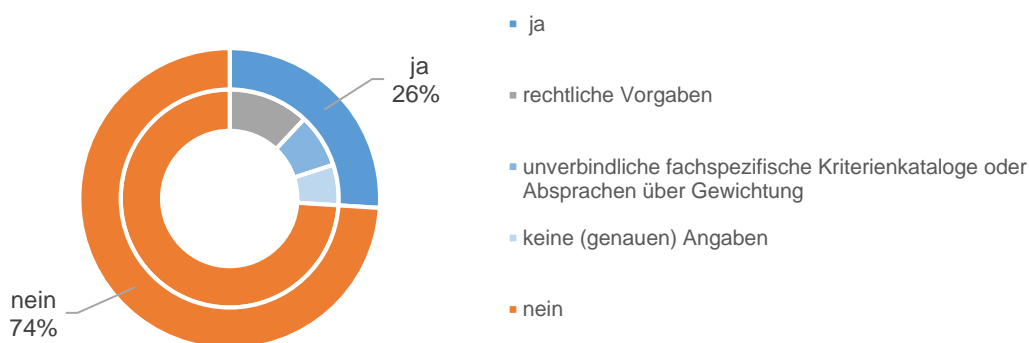


Abbildung 13: Anteil an Lehrkräften mit Absprachen im Kollegium

4.3.1.7 Bezugsnormorientierung

Zu welcher Bezugsnorm die Lehrkräfte neigen (Forschungsfrage 1.9), um die mündliche Mitarbeit zu bewerten, wurde mithilfe einer Ratingskala untersucht. Die Frage nach dem festgelegten Erwartungshorizont bezieht sich auf die kriteriale Bezugsnorm, der Vergleich der Leistung mit den anderen aus der Lerngruppe auf die soziale Bezugsnorm und die Bewertung anhand des individuellen Lernfortschritts auf die individuelle Bezugsnorm. Dabei wählten die Lehrenden aus, inwiefern die jeweilige Bezugsnorm für ihre Bewertung der mündlichen Mitarbeit angewandt wird.

¹ Dabei wurden meist die hessischen Vorgaben der Gewichtung der mündlichen Note im Vergleich zu den schriftlichen Leistungen für ein spezielles Fach erläutert.

Für $n_{gesamt} = 50$ ergaben sich folgende Mittelwerte, die in Abbildung 14 dargestellt sind. Die Standardabweichung wird durch die schwarzen Balken repräsentiert.

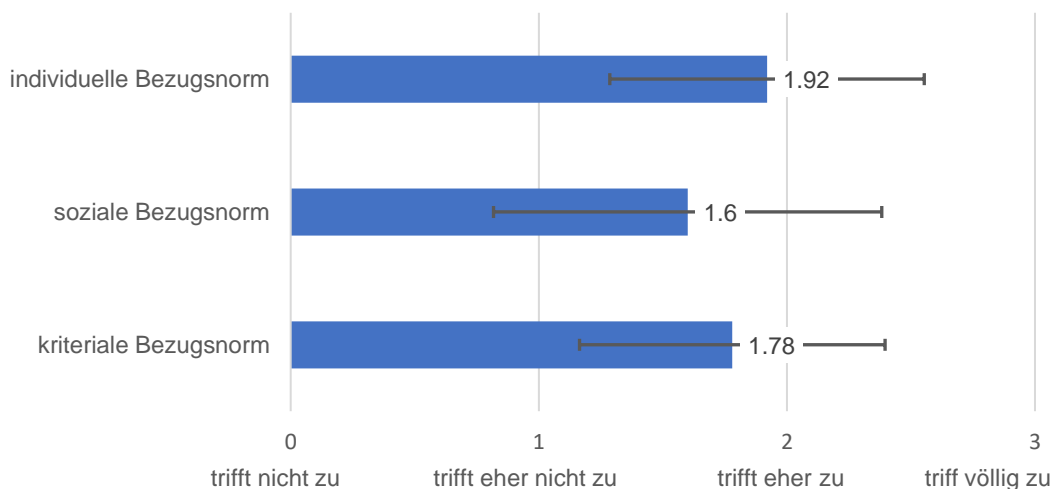


Abbildung 14: Verteilung der Bezugsnormorientierung der gesamten Stichprobe

Hierbei lässt sich feststellen, dass die Lehrkräfte bei der Bewertung mündlicher Mitarbeit alle drei Bezugsnormen anwenden, die dabei ähnlich bedeutsam sind. Die individuelle Bezugsnorm ist bei der gesamten Stichprobe der wichtigste Bezugsrahmen und die soziale Bezugsnorm am wenigsten bedeutsam. Die Streuung ist bei der sozialen Bezugsnorm (0.78) etwas größer als bei der kriterialen (0.62) und der individuellen Bezugsnorm (0.63).

In Abbildung 15 wird außerdem zwischen Lehrkräften mit wenig, mittlerer und hoher Berufserfahrung unterschieden. In den darauf folgenden drei Tabellen werden die Effekte bei den Abweichungen dargestellt.

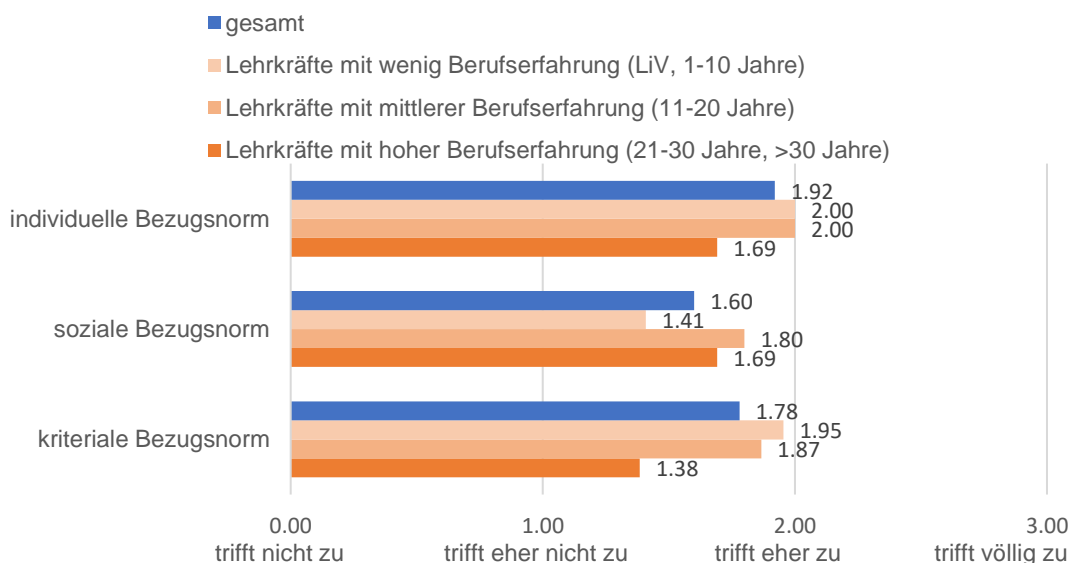


Abbildung 15: Verteilung der Bezugsnormorientierung nach Berufserfahrung

Tabelle 21: Abweichung der Gewichtung der individuellen Bezugsnorm nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	2.00	0.53	0.132
Gesamte Stichprobe	50	1.92	0.63	
Mittlere Berufserfahrung	15	2.00	0.65	0.125
Gesamte Stichprobe	50	1.92	0.63	
Hohe Berufserfahrung	13	1.69	0.75	0.346
Gesamte Stichprobe	50	1.92	0.63	

Tabelle 22: Abweichung der Gewichtung der sozialen Bezugsnorm nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	1.41	0.91	0.232
Gesamte Stichprobe	50	1.60	0.78	
Mittlere Berufserfahrung	15	1.80	0.56	0.271
Gesamte Stichprobe	50	1.60	0.78	
Hohe Berufserfahrung	13	1.69	0.75	0.119
Gesamte Stichprobe	50	1.60	0.78	

Tabelle 23: Abweichung der Gewichtung der kriterialen Bezugsnorm nach Berufserfahrung

Gruppe	Gruppen- größe n	Mittelwert	Standard- abweichung	Effekt- größe d
Wenig Berufserfahrung	22	1.95	0.65	0.278
Gesamte Stichprobe	50	1.78	0.62	
Mittlere Berufserfahrung	15	1.87	0.52	0.146
Gesamte Stichprobe	50	1.78	0.62	
Hohe Berufserfahrung	13	1.38	0.51	0.664
Gesamte Stichprobe	50	1.78	0.62	

Hierbei ist mit einem kleinen Effekt erkennbar, dass Lehrpersonen mit hoher Berufserfahrung die individuelle Bezugsnorm weniger anwenden als die gesamte Stichprobe (siehe Tabelle 21).

Aus Tabelle 23 wird ersichtlich, dass diese Gruppe sich für die Bewertung mündlicher Mitarbeit ebenfalls weniger an der kriterialen Bezugsnorm, mit einem mittleren Effekt, orientiert als die gesamte Stichprobe im Mittel.

Bei der kriterialen Bezugsnorm lässt sich dagegen eine kleine gegensätzliche Tendenz bei Lehrkräften mit wenig Berufserfahrung feststellen, die ihre Bewertung eher danach ausrichten als bei der gesamten Stichprobe. Des Weiteren schätzen die Lehrkräfte mit wenig Berufserfahrung die soziale Bezugsnorm, mit einem kleinen Effekt, als weniger relevant ein, wie in Tabelle 22 erkennbar wird.

4.3.2 Dokumentation

Der größte Teil der befragten Lehrpersonen (86%) gab an, die mündliche Mitarbeit mit einer Tabelle zu dokumentieren, wobei etwa zu gleichen Teilen Skalen mit Noten oder --/o/+ /++ sowie Kurzbeschreibungen (mit Auffälligkeiten) eingetragen werden. 14 % der Lehrkräfte gaben an, keine Tabelle zu verwenden. Sie benannten folgende Methoden, um die mündliche Mitarbeit zu dokumentieren: Die Bewertung jedes Lernenden anhand eines Kriterienkatalogs unter Verwendung der Skala --/o/+ /++ oder Kurzbeschreibungen, wo überwiegend Auffälligkeiten notiert werden. Außerdem wurde angegeben, dass einzelne Lernende gezielt beobachtet werden, wenn sie zuvor durch wenig Mitarbeit kaum aufgefallen sind.

Die Antworten in Bezug auf die Regelmäßigkeit der Dokumentation sind in Abbildung 16 dargestellt.

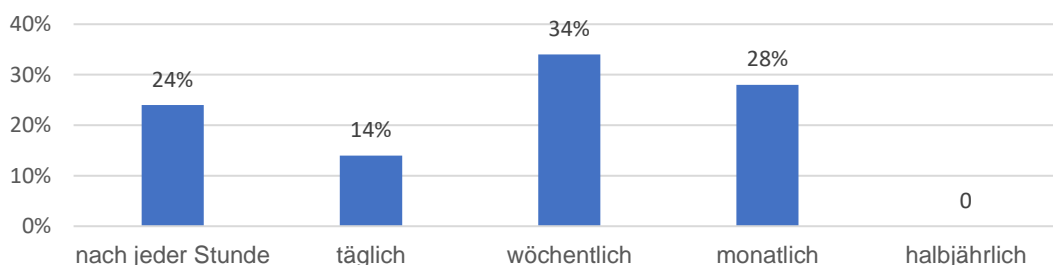


Abbildung 16: Regelmäßigkeit der Dokumentation der gesamten Stichprobe

Die Zeitspanne der Dokumentation reicht von nach jeder Unterrichtsstunde bis zu monatlich. Der mit 34% größte Anteil der Lehrkräfte hält die Leistungen der mündlichen Mitarbeit wöchentlich fest. Danach folgen monatlich (28%) und nach

jeder Stunde (24%). Täglich werden die Leistungen von 14% der Lehrenden festgehalten und keine Lehrperson gab an, sich halbjährlich Noten zu notieren¹.

Die Abbildung 17 zeigt, wie die einzelnen Gruppen mit verschieden ausgeprägter Berufserfahrung diese Frage beantwortet haben:

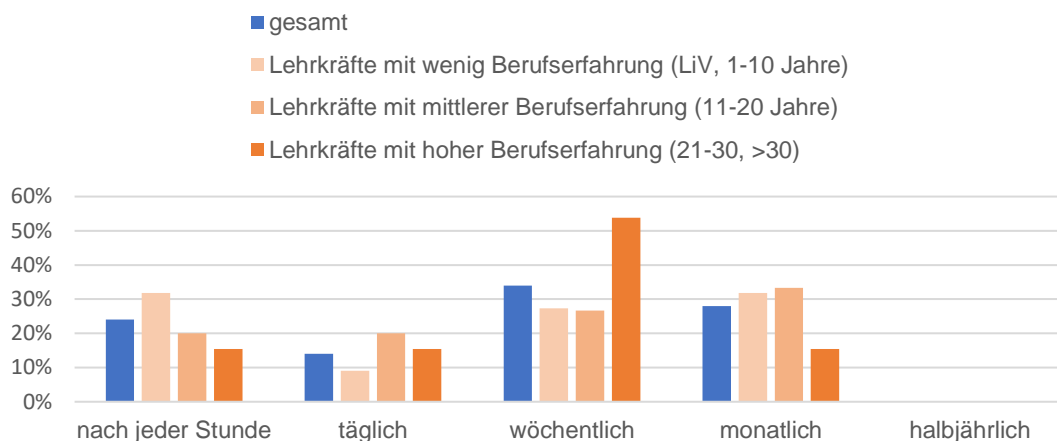


Abbildung 17: Regelmäßigkeit der Dokumentation nach Berufserfahrung

Die prozentualen Angaben beziehen sich dabei auf 100% der Antworten aus einer Berufserfahrungs-Gruppe. Dabei ist auffällig, dass eher Lehrende mit weniger Berufserfahrung nach jeder Stunde dokumentieren als solche mit mittlerer oder hoher Berufserfahrung. Über 53% der Lehrpersonen mit hoher Berufserfahrung dokumentieren die mündliche Mitarbeit wöchentlich.

Weiterhin wurden die Lehrenden befragt, ob sie mit der Häufigkeit der Dokumentation einverstanden sind. Im Ergebnis war die Mehrheit der Lehrkräfte (68%) mit der Häufigkeit der Dokumentation zufrieden. Die übrigen Lehrenden (32%) wünschten sich, häufiger dokumentieren zu können. Die von diesen Teilnehmenden angegebenen Gründe für eine seltenere als gewünschte Dokumentation der mündlichen Mitarbeit werden in Abbildung 18 dargestellt. Dabei gaben einzelne Lehrpersonen auch mehrere Gründe an.

¹ Halbjährlich sind die mündlichen Noten für die Zeugnisnoten fällig.

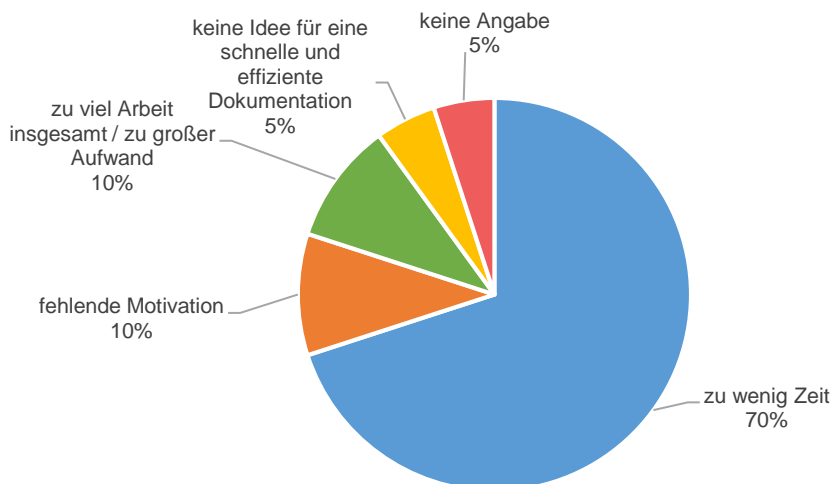


Abbildung 18: Genannte Gründe für eine seltenere als gewünschte Dokumentation

Der mit 70% am meisten genannte Grund ist mangelnde Zeit im Unterrichtsalltag. Außerdem wurde mit 10% fehlende Motivation und mit ebenfalls 10% der zu große Aufwand als Grund benannt. Des Weiteren fehlen Ideen für eine schnelle und effizientere Dokumentation.

4.3.3 Reflexion

Die Lehrenden konnten mithilfe einer Lickert-Skala angeben, wie sie ihre eigene Bewertung beurteilen. Bei dem ersten Merkmal sollte eingeschätzt werden, inwiefern der Bewertungsprozess unabhängig von der Lehrperson erfolgt (Objektivität). Für wie zuverlässig die Bewertung mündlicher Mitarbeit gehalten wird, sodass auch Wiederholungen zu demselben Ergebnis führen würden (Reliabilität), wurde ebenfalls abgefragt. Des Weiteren wurden die Lehrenden nach der Gültigkeit gefragt, dass ihre Bewertung auch die der tatsächlichen Leistungen der mündlichen Mitarbeit widerspiegelt (Validität). Weiterhin wurde ermittelt, für wie transparent die Lehrkräfte ihre Bewertungskriterien für SuS und Eltern halten. Außerdem wurden sie danach gefragt, inwiefern die SuS mit ihrer Bewertung einverstanden sind (Akzeptanz). Die Ergebnisse für $n_{gesamt} = 50$ sind in Abbildung 19 veranschaulicht:

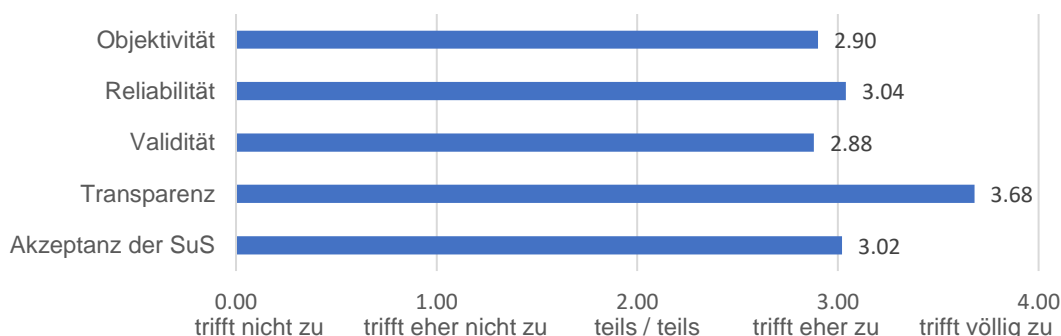


Abbildung 19: Reflexion der Bewertung mündlicher Mitarbeit

Die Mittelwerte zeigen, dass die Lehrenden den einzelnen Aussagen bezüglich Objektivität, Reliabilität, Validität und Akzeptanz eher zustimmen. Bei dem Merkmal der Transparenz stimmten die Lehrkräfte sogar fast völlig zu.

Auf die Darstellung der Antworten, separat nach Berufserfahrung, wurde an dieser Stelle bewusst verzichtet, da keine Abweichungen erkennbar waren.

Bei der offenen und freiwilligen Frage am Ende bekamen die Lehrenden die Gelegenheit, mögliche Schwierigkeiten bei der Bewertung mündlicher Mitarbeit zu benennen. Dazu haben sich insgesamt zehn Lehrkräfte geäußert. Einzelne Aussagen sollen hier zitiert werden:

„Mündliche Notengebung ist für mich sehr schwierig. Manchmal wünsche ich mir mehr Zeit zu haben darüber nachzudenken und einzelne Wortmeldungen wörtlich zu notieren.“

„Solange der Notenfindungsprozess für die Schüler transparent ist gibt es aus meiner Sicht keine Probleme“.

„Es nervt mich, ständig im Bewertungsmodus zu sein. Aber ohne diesen ist die Arbeitseinstellung der meisten Schüler leider nicht ausreichend.“

„Eine Schwierigkeit, die ich sehe, ist, dass es mir [...] schwerer fällt, einem Schüler, der nur schwache schriftliche Leistungen erbringt, in der mündlichen Mitarbeit eine gute oder sehr gute Note zu geben. Im Gegensatz dazu profitieren Schüler, die schriftlich gut oder sehr gut sind, von ihrer schriftlichen Leistung, da ich dazu neige, sie vor diesem Hintergrund mündlich ‚wohlwollender‘ zu bewerten (davon profitieren die typischen ‚stillen Mädchen‘, die schriftlich leistungsstark sind)“.

„Natürlich ist die Bewertung mündlicher Leistungen immer eine Mischung aus wenigen objektiven und vielen subjektiven Aspekten. Ich hoffe, diese mit einem gewissen Grad an Erfahrung, Schülerzugewandtheit und Fairness zu betrachten. Um meine Einschätzung zu vervollständigen, lasse ich in der Mittel- und Oberstufe Selbsteinschätzungsbögen [sic] ausfüllen oder begründete Selbsteinschätzungen von den [sic] Schüler schreiben. Diese decken sich in der Regel zu 95% mit meiner Wahrnehmung, die Fälle, in denen die Wahrnehmung stark differiert, betrachte ich mir näher und bespreche sie mit/ begründe sie gegenüber den Schülern.“

In diesen ausgewählten Zitaten wird zum einen der Faktor Zeit als belastend genannt. Außerdem wurden die Hemmungen angesprochen, die mündliche Mitarbeit weit entfernt von den Ergebnissen der schriftlichen Leistungen zu

bewerten. Eine Lehrperson beklagt sich über die Rolle der ständigen Bewertungsperson im Unterricht („Bewertungsmodus“). Es werden jedoch auch Lösungen benannt, womit die Akzeptanz über die Bewertung erhöht werden kann. Dabei spielt zum einen die Transparenz der Kriterien und zum anderen das Einbeziehen der Lernenden in den Bewertungsprozess eine Rolle.

4.4 Diskussion

Diese Studie beschäftigte sich mit der Praxis der Bewertung mündlicher Mitarbeit aus der Perspektive der Lehrperson. Sie sollte also die persönlichen Vorgehensweisen und Einstellungen der Lehrkräfte erheben.

4.4.1 Zusammenfassung der Ergebnisse

Zunächst wird eine Irritation rund um die Begriffe „mündliche Mitarbeit“ und „mündliche Note“ deutlich. Obwohl nach Kriterien für die mündliche Mitarbeit gefragt wurde, wurden mit fast einem Viertel der Nennungen Bestandteile der mündlichen Note angegeben, die vor allem schriftliche Leistungen enthielten. Bei der Frage nach schulischen Vorgaben und Absprachen im Kollegium wurden weiterhin häufig rechtliche Vorgaben benannt. Ansonsten konnten nur marginal Hilfestellungen oder Absprachen in der Schule oder im Kollegium festgestellt werden. Die Bewertung mündlicher Mitarbeit liegt damit tatsächlich eher in der individuellen Verantwortung der einzelnen Lehrperson (Forschungsfrage 1.8). Die in der Studie von Kirk (2004) konstatierte Uneinheitlichkeit und Unsicherheit können damit in dieser Studie bestätigt werden.

Sowohl die Ergebnisse der offenen Frage nach den Kriterien als auch in der Rankingfrage zeigen die Dominanz der Kriterien Qualität und Quantität (Forschungsfrage 1.1 und 1.2). Die Praxis zeigt damit, dass auch quantitative Aspekte, wie die Häufigkeit und Regelmäßigkeit der mündlichen Mitarbeit, einen großen Einfluss auf die Bewertung haben, wie schon Krieger (2003) festgestellt hat.

In der Studie dieser Arbeit wurde das erste Mal gezielt nach der *Berücksichtigung der Extraversion* gefragt, um zu erfahren, ob sozial gehemmte Lernende auch mit einer weniger häufigen, qualitativ aber sehr guten Beteiligung ebenfalls eine gute Bewertung erzielen können. Die Ergebnisse zeigten eine eher geringe Berücksichtigung der verschieden ausgeprägten Extraversion (Forschungsfrage 1.3).

Die *Gesprächsfähigkeit* wurde zwar in der offenen Frage mit 7% eher selten benannt, bei der Gewichtung der Kriterien wurde sie jedoch direkt nach den Aspekten der Qualität und Quantität eingestuft. Die Fähigkeiten des Zuhörens, Argumentierens und Moderierens zählen damit eher zu den wichtigeren Kriterien bei der Bewertung mündlicher Mitarbeit (Forschungsfrage 1.4). Auffallend zeigte sich in der Auswertung der Rangfolge nach Aufgabenfeldern eine leicht erhöhte Beachtung der Gesprächsfähigkeit in Fächern des Aufgabenfeldes 2 (Forschungsfrage 1.6), die in weiterführenden Studien überprüft werden könnte.

Die in Abschnitt 3.3.3.2 beschriebene Herausforderung der sachfremden Beeinflussung durch schriftliche Leistungen zeigte sich in der Bemerkung einer Lehrperson in der Frage nach Schwierigkeiten bei der Bewertung. Die schriftlichen Leistungen (meist Noten in Klassenarbeiten) werden zwar nicht direkt als Kriterium für die Bewertung angeführt, in dem Rankingplatz (7/9) wird dennoch eine, wenn auch geringere, Beachtung deutlich (Forschungsfrage 1.5). Hierbei gab es Auffälligkeiten in Bezug auf eine erhöhte Berücksichtigung in den Fächern des Aufgabenfeldes 3 und eine geringere Beachtung im Aufgabenfeld 4 (Forschungsfrage 1.6). Die Erledigung von Hausaufgaben, die nach der Definition kein Teil der mündlichen Mitarbeit ist, wird mit einem Ranking von (6/9) etwas wichtiger eingeschätzt. Dies könnte mit der oben beschriebenen Irritation über die Begriffe zusammenhängen.

Das *Sozial-/ Störverhalten* war das im Durchschnitt am wenigsten bedeutsam eingestufte Kriterium. Die Antworten des Aufgabenfeldes 4 (Sport) zeigten hierbei eine stark erhöhte Relevanz im Vergleich zum Mittelwert aller Aufgabenfelder (Forschungsfrage 1.6).

Insgesamt waren bei dem Stellenwert der Kriterien eher größere Unterschiede zwischen den Aufgabenfeldern als zwischen den Gruppen nach Berufserfahrung erkennbar. Hierbei zeigten sich aber immerhin bei Lehrkräften mit viel Berufserfahrung, mit eher kleinen Effekten, eine erhöhte Relevanz der Kriterien *Qualität der Beiträge* und *Erledigung der Hausaufgaben* sowie einer geringeren Relevanz der *Berücksichtigung der (Lern-) Entwicklung* der SuS. Für Lehrende mit wenig Berufserfahrung konnte eine etwas erhöhte *Berücksichtigung des Sozial-/ Störverhaltens* konstatiert werden (Forschungsfrage 1.7).

Die befragten Lehrenden gaben außerdem an, alle drei Bezugsnormen (ähnlich stark ausgeprägt) als Bezugsrahmen für die Bewertung mündlicher Mitarbeit zu nutzen, obwohl insbesondere die soziale Bezugsnorm für die schulische Leistungsbewertung in der Literatur kritisiert wird, wie in Abschnitt 3.2.2.1

Bezugsnormen dargestellt wurde. Dabei tendierten in der Stichprobe dieser Studie die Lehrkräfte mit wenig Berufserfahrung eher zur kriterialen und individuellen Bezugsnorm und eher weniger zur sozialen Bezugsnorm. Solche mit mittlerer Berufserfahrung neigten eher zur individuellen und sozialen Bezugsnorm (jeweils kleine Effekte). Lehrende mit viel Berufserfahrung wichen insofern vom Durchschnitt ab, dass sie weniger die individuelle (kleiner Effekt) und kriteriale (mittlerer Effekt) Bezugsnorm nutzen, als der Durchschnitt der gesamten Stichprobe (Forschungsfrage 1.9).

Die Dokumentation mündlicher Mitarbeit geschieht überwiegend in tabellarischer Form, mit Noten oder kleinen Bemerkungen, wie auch in der Befragung von Kirk (2004) ermittelt wurde (Forschungsfrage 2.1). Hinsichtlich der Regelmäßigkeit konnten ähnliche Ergebnisse wie bei Krieger (2003) und Kirk (2004) festgestellt werden: Die Lehrkräfte dokumentieren die Leistungen in unterschiedlich großen Abständen, nach jeder Stunde bis monatlich. In dieser Befragung gab die Mehrheit der Teilnehmenden an, wöchentlich zu dokumentieren (Forschungsfrage 2.2). Weiterhin gab es kleine Unterschiede zwischen den Lehrergruppen nach Berufserfahrung insofern, dass Lehrkräfte mit wenig Berufserfahrung eher nach jeder Stunde Notizen anfertigen als die anderen Gruppen, wohingegen die Mehrheit (53%) der Lehrenden mit viel Berufserfahrung wöchentlich dokumentieren (Forschungsfrage 2.3).

Neu war außerdem die Frage nach der Zufriedenheit über die Zeitspanne der Dokumentation. Immerhin knapp ein Drittel der Lehrenden wünschte sich eine Veränderung und beklagte sich vorwiegend über mangelnde Zeit im Unterrichtsalltag.

In der Reflexion über die eigene Bewertungspraxis (Forschungsfrage 3.2) zeigte sich, dass die Lehrkräfte ihre Bewertung als tendenziell unproblematisch ansehen und sogar eher für objektiv, reliabel und valide halten. Besonders selbstbewusst antworteten sie bezüglich der Transparenz der Kriterien. Bei der Akzeptanz bestätigt sich der Eindruck der früheren Studie von Krieger (2003), wo die Lehrkräfte die Bewertung mündlicher Mitarbeit als eher reibungslosen Vorgang bezeichneten und die Akzeptanz der SuS einschätzen, obwohl die dort ebenfalls befragten ehemaligen SuS durchaus auch Probleme benannten.

Die Bemerkungen der Lehrkräfte (siehe Zitate) waren allerdings auch dadurch gekennzeichnet, dass die befragten LuL den eigenen Bewertungsprozess durchaus kritisch reflektierten und bestimmte Urteilstendenzen erkannten (Forschungsfrage 3.1). Außerdem wurde angegeben, dass teilweise die SuS durch Selbsteinschätzungsbögen in den Vorgang miteinbezogen werden.

4.4.2 Limitation der Ergebnisse und des Forschungsdesigns

Rückschlüsse auf die Grundgesamtheit der hessischen Gymnasiallehrerinnen und -lehrer können aber nur mit Vorsicht gezogen werden, da die Repräsentativität durch die Größe und Akquirierung der Stichprobe nicht gewährleistet sein kann. Dies sollte insbesondere bei der sehr kleinen Teilgruppe der Sportlehrerinnen und -lehrer (n=5) berücksichtigt werden.

Mit dem Querschnittsdesign handelt es sich weiterhin lediglich um eine Momentaufnahme der aktuellen Praxis. Außerdem muss die Aussagekraft der Fragen über die Kriterien, inklusive der Vorgaben und Absprachen, eingeschränkt werden, da Begriffsunklarheiten und Unsicherheiten festzustellen waren. Der Vergleich einzelner Teilgruppen der gesamten Stichprobe ergab meist kleine bis mittel effektive Unterschiede, die Gegenstand weiterer Untersuchungen sein könnten, um die Ergebnisse entweder zu bestätigen oder zu widerlegen.

Des Weiteren kann diese Befragung nicht die „Wirklichkeit“ der Bewertungspraxis abbilden, da sie lediglich die Aussagen der Lehrkräfte erheben konnte. Unklar ist zudem, ob die Teilnehmenden allgemein eher anspruchs- oder wirklichkeitskonform geantwortet haben. Außerdem bleibt ungeklärt, ob es bestimmte Antwortverzerrungen, insbesondere durch soziale Erwünschtheit oder eine Zustimmungstendenz, gegeben hat.

Insgesamt konnte diese Studie durchaus einen spannenden Einblick in die Praxis der Bewertung mündlicher Mitarbeit, aus Lehrerperspektive, gewähren. Einige Ergebnisse aus früheren Studien konnten bestätigt werden. Dieses Feld der mündlichen Leistungsbewertung bedarf gewiss noch weiterer, differenzierter Untersuchungen.

5. Fazit

Diese Arbeit widmete sich der Bewertung mündlicher Mitarbeit, da diese in der bisherigen Forschung trotz des hohen Stellenwerts für die schulische Leistungsbewertung stark vernachlässigt wurde. Mithilfe der eigenen Studie wurde versucht, einen Einblick in die Bewertungspraxis, aus dem Blickwinkel der Lehrkräfte, zu erhalten und einen Ausblick auf weitere Forschungsperspektiven zu eröffnen.

Aufgrund der messtheoretischen Qualitätsanforderungen wird ein hoher *Anspruch* an schulische Leistungsbewertung deutlich: Bei der Feststellung der mündlichen

Mitarbeit soll dem Grundsatz der Gütekriterien Objektivität, Reliabilität und Validität gefolgt werden. Bei der Leistungsbeurteilung und -bewertung sollen soziale Vergleiche vermieden und dafür überwiegend ein kriterialer Bezugsrahmen, ergänzend auch die individuelle Bezugsnorm, angewandt werden.

Nach den rechtlichen Vorgaben des Landes Hessen soll zudem die individuelle Lernentwicklung und Leistungsbereitschaft des Lernenden berücksichtigt werden. Außerdem sind Lehrkräfte verpflichtet, die Lernenden und Eltern über Bewertungskriterien und Leistungsstand regelmäßig zu informieren. Die Bewertungsgegenstände der mündlichen Mitarbeit sind in den hessischen Vorgaben nicht enthalten und deren Auswahl und Gewichtung liegen daher im persönlichen Ermessen jeder Lehrperson.

Eine *Chance* der kontinuierlichen Bewertung mündlicher Mitarbeit besteht vor allem darin, dass Leistungen ganzheitlich und über einen längeren Zeitraum festgestellt werden können, im Gegensatz zu eher punktuellen Klassenarbeiten. Davon profitieren insbesondere solche SuS, die entweder unter erhöhter Prüfungsangst leiden oder gerne kreative Beiträge leisten, die in einer standardisierten Klassenarbeit weniger zur Geltung kommen.

Dennoch gibt es zahlreiche *Herausforderungen* bei der Bewertung mündlicher Mitarbeit, die es zu meistern gilt:

Die Schwierigkeit besteht für eine Lehrperson vor allem darin, während des Unterrichtsgeschehens als beteiligter Akteur, quasi „nebenbei“, die SuS zu beobachten und dabei zu gewährleisten, dass auch möglichst alle Lernenden die Gelegenheit erhalten, einen Wortbeitrag zu leisten. Des Weiteren wirken sich, neben dem Aufrufverhalten der Lehrkraft, auch die unterschiedliche Extraversion der Lernenden sowie die situativen Voraussetzungen auf das Meldeverhalten der SuS aus.

Die „subjektiven“ Verfahren zur Leistungsfeststellung bieten, insbesondere bei mündlichen Leistungen, zu wenige Möglichkeiten zur Vereinheitlichung und Schematisierung, um den hohen Anforderungen überhaupt gerecht zu werden.

Die Bewertung mündlicher Mitarbeit basiert meist auf unsystematischen Gelegenheitsbeobachtungen, bei der eine Lehrperson wegen der Komplexität des Kontextes nicht alle relevanten Aspekte wahrnehmen kann. Dies hängt u. a. auch von kognitiven Faktoren, wie Konzentration und Erinnerungsvermögen ab. Durch die Flüchtigkeit der mündlichen Wortbeiträge zum Unterrichtsgespräch können bis zum Zeitpunkt der Dokumentation Erinnerungslücken oder Verzerrungen entstehen. Dennoch gibt es in den rechtlichen Vorgaben in Hessen keinerlei Verpflichtungen zur regelmäßigen Dokumentation. Mit Gedächtnisprotokollen oder

systematischen Beobachtungsverfahren könnte der Einschränkung, dass Leistungen nicht frei von subjektiven Einflüssen beobachtet werden können, entgegen gewirkt werden. Um mögliche Urteilsfehler zu vermeiden, sollten Lehrende stets ihre eigene Bewertungspraxis reflektieren, um persönliche Tendenzen zu erkennen. Des Weiteren stellt sich die Frage, ob und wie mit weniger offensiven, introvertierten SuS umgegangen werden kann, um sie nicht unbeabsichtigt systematisch zu benachteiligen.

Schließlich besteht eine grundsätzliche Schwierigkeit darin, eine klare Trennung zwischen den widersprüchlichen Bedingungen von Lern- und Prüfungssituationen zu erreichen. Bei der Konzentration auf eine möglichst „gerechte“ und kontinuierliche Bewertung sollte nicht die Bedeutung des Lernens in einer lernförderlichen und angstfreien Umgebung vernachlässigt werden.

Insgesamt stellt sich die Bewertung mündlicher Mitarbeit als vielschichtiges Problemfeld dar.

Daher war es spannend in einer eigenen Studie, zu untersuchen, wie die *Wirklichkeit* der Bewertungspraxis von Lehrkräften selbst wahrgenommen wird.

In den Ergebnissen dieser Studie waren bei den Begriffen der mündlichen Mitarbeit und mündlichen Note sowie den schulischen Vorgaben und kollegialen Absprachen Irritationen festzustellen, die mit fehlenden bzw. vagen Formulierungen in den Vorgaben des Landes Hessen zusammenhängen könnten. Damit zeigen sich diesbezüglich ähnliche Unsicherheiten auf Seiten der Lehrperson, wie bei der Umfrage von Kirk (2004) schon bemerkt wurden.

Die Lehrkräfte der Studie dieser Arbeit gaben außerdem an, nur vereinzelt Vorgaben auf Schulebene oder Hilfestellungen im Kollegium zu erhalten. Damit liegen die Suche nach geeigneten Kriterien und die Durchführung der Bewertung mündlicher Mitarbeit für den größten Teil der Lehrkräfte vollständig in der persönlichen Verantwortung.

Wie durch Krieger (2003) bereits erhoben wurde, bestätigten sich die Qualität und Quantität als wichtigste Kriterien für die Bewertung mündlicher Mitarbeit. Dadurch, dass die Häufigkeit der Beiträge derart relevant für die Bewertung zu sein scheint, stellt sich die Frage, warum die Unterschiedlichkeit der Lernenden bezüglich Extraversion vergleichsweise wenig berücksichtigt wird.

Die Gesprächsfähigkeit zeigte sich weiterhin als tendenziell wichtiges Kriterium zur Bewertung mündlicher Mitarbeit, insbesondere für Fächer, wie Geschichte, Politik und Wirtschaft, Ethik, Religion und Geografie.

Die Wirkung von Vorinformationen, in Bezug auf schriftliche Leistungen in Klassenarbeiten, wurde von den befragten Lehrpersonen als eher weniger relevant

eingeschätzt. Dennoch sollte dieser Faktor der möglichen Verzerrungen, die besonders in Fächern wie Mathematik, Physik, Biologie, Informatik etc. und am wenigsten in Sport festzustellen waren, nicht unterschätzt und durch zukünftige Forschung überprüft werden. Denn die Beeinflussung durch sachfremde Aspekte geht zu Lasten der Gültigkeit der Bewertung mündlicher Mitarbeit.

Die in der Literatur empfohlene vorrangige Anwendung der kriterialen Bezugsnorm wird in der von den Lehrkräften angegebenen Wirklichkeit nicht umgesetzt. Stattdessen zeigt sich in dieser Studie eine in etwa ähnliche Ausprägung aller drei Bezugsnormen. Die kriteriale und individuelle Bezugsnorm spielen besonders für Lehrkräfte mit wenig Berufserfahrung eine, im Vergleich zum Durchschnitt, erhöhte und bei solchen mit viel Berufserfahrung eher eine geringere Rolle. Diese Ergebnisse könnten auf den unterschiedlich großen zeitlichen Abstand zur Lehrerausbildung basieren und sollten, durch weiterführende Studien, überprüft werden.

Durch die Flüchtigkeit mündlicher Mitarbeit besteht durch eine natürliche Begrenztheit der Wahrnehmung und des Erinnerungsvermögens ein erhöhter Bedarf an regelmäßiger Dokumentation, um möglichen Verzerrungen entgegen wirken zu können. Die für diese Arbeit befragten Lehrkräfte halten die mündliche Mitarbeit mehrheitlich wöchentlich und in einer Tabelle fest. Insgesamt reicht die Spanne jedoch von sehr kleinen Abständen (nach jeder Stunde) bis zu eher großen (monatlich). Daraus entsteht die Frage nach den, für den hohen Anspruch und die vielen zu beobachtenden Lernenden, angemessenen Zeitabständen der Dokumentation und wie Lehrende diese, vor dem Hintergrund des beklagten Zeitmangels, umsetzen können.

In der Studie von Krieger (2003) sollten Lehrkräfte bereits ihre Bewertungspraxis reflektieren. Dabei kristallisierte sich eine Diskrepanz zwischen Schüler- und Lehrerwahrnehmung hinsichtlich der Wahrnehmung der Reibungslosigkeit der Bewertung heraus. In der Studie dieser Arbeit sollten die Lehrkräfte, zusätzlich zur Akzeptanz und Transparenz, auch die Objektivität, Reliabilität und Validität einschätzen. Die Ergebnisse zeigen, dass die Angaben der Lehrkräfte ein eher positives Bild zeichnen, was anhand der oben beschriebenen Anforderungen und Schwierigkeiten widersprüchlich erscheint. Besonders gut schätzten sich die befragten Lehrkräfte in Bezug auf die Transparenz der Kriterien ein, wohingegen bei Krieger (2003) im Vergleich dazu nur 54% der LuL davon überzeugt waren, dass ihre Kriterien transparent für die SuS sind. Lediglich vereinzelt äußerten sich Lehrkräfte kritisch reflektiert über Schwierigkeiten, auch hinsichtlich unbeabsichtigter Beeinflussung durch beispielsweise schriftliche Leistungen.

Schließlich wurde auch die Möglichkeit der Einbeziehung der Lernenden in den Vorgang beschrieben, um so die Transparenz und Akzeptanz erhöhen zu können.

Auch wenn diese Studie die Wirklichkeit lediglich aus der Perspektive der befragten Lehrpersonen und mit Einschränkungen der Repräsentativität abbilden kann, gestattet sie damit einen wichtigen Einblick in die praktische Umsetzung der Bewertung mündlicher Mitarbeit.

Insgesamt kann eine große Diskrepanz zwischen dem hohen *Anspruch* und der *Wirklichkeit* der alltäglichen Bewertungspraxis konstatiert werden. Zum einen scheint eine stärkere Sensibilisierung der Lehrkräfte für die zahlreichen und vielseitigen Herausforderungen, gerade durch die individuelle Verantwortung, eine elementare Grundlage für zukünftige Weiterentwicklungen zu sein. Zum anderen besteht, aufgrund des Mangels konkreter rechtlicher Vorgaben, die Notwendigkeit eines Konsenses über geeignete Kriterien und Hilfestellungen bzw. Ideen zu realisierbaren Beobachtungs- und Dokumentationsmethoden auf Schulebene oder im Kollegium. Damit könnten Lehrende bei der Bewertung mündlicher Mitarbeit entlastet und unterstützt werden. Dies könnte zu einer Vereinheitlichung und damit zur Erleichterung der Legitimation gegenüber Lernenden und ihren Eltern sowie zu einer Verringerung der Unsicherheiten bei allen Beteiligten beitragen.

6. Literaturverzeichnis

- Abraham, U. (2012). *Sprechen als reflexive Praxis. Mündlicher Sprachgebrauch in einem kompetenzorientierten Deutschunterricht*. 1. Aufl. Stuttgart: Fillibach bei Klett.
- Arnold, K. H. & Jürgens, E. (2001). *Schülerbeurteilung ohne Zensuren* (Vol. 8). Neuwied [u. a.]: Luchterhand.
- Avenarius, H. & Füssel, H.-P. (2010). *Schulrecht: Ein Handbuch für Praxis, Rechtsprechung und Wissenschaft*. Kronach: Link.
- Becker-Mrotzek, M. & Brünner, G. (Hrsg.) (2004). *Analyse und Vermittlung von Gesprächskompetenz* (Vol. 43). Gesellschaft für Angewandte Linguistik; Jahrestagung der Gesellschaft für Angewandte Linguistik (GAL). Frankfurt am Main [u. a.]: Lang.
- Beckmann, H.-K. (Hrsg.) (1978). *Leistung in der Schule* (Vol. 166). Braunschweig: Westermann.
- Berkemeier, A. & Spiegel, C. (2014). In der Schule Gesprächsfähigkeit fordern und fördern: Moderieren – Argumentieren – Zuhören. In: Grundler, E. & Spiegel, C. (Hrsg.) *Konzeptionen des Mündlichen. Wissenschaftliche Perspektiven und didaktische Konsequenzen. Symposium Deutschdidaktik (Mündlichkeit, Bd. 3)*. Bern: Hep der Bildungsverl. S. 120-142.
- Birkel, P. (1984). Beurteilung mündlicher Prüfungsleistungen. In: Heller, K. A. (Hrsg.) *Leistungsdiagnostik in der Schule. 4., völlig neubearbeitete Auflage*. Bern/ Stuttgart/ Toronto: Huber. S. 229-236.
- Bittner, S. (2006). *Das Unterrichtsgespräch: Formen und Verfahren des dialogischen Lehrens und Lernens*. Bad Heilbrunn: Klinkhardt.
- Bodenmann, G. (2006). Beobachtungsmethoden. In: Petermann, F. (Hrsg.) *Handbuch der psychologischen Diagnostik* (Vol. 4). Göttingen [u. a.]: Hogrefe. S. 151-159.
- Boer, H. d. & Reh, S. (2012). *Beobachtung in der Schule - Beobachten lernen*. Wiesbaden: Springer VS.
- Böhm, T. (2008). *Zentrale Fragen zur Leistungsbeurteilung und zum Prüfungsrecht* (Vol. 3). Kronach: Luchterhand.
- Bortz, J. (1984). *Lehrbuch der empirischen Forschung: Für Sozialwissenschaftler*. Berlin [u. a.]: Springer.
- Breidenstein, G. (2006). *Teilnahme am Unterricht: Ethnographische Studien zum Schülerjob* (Vol. 24). Wiesbaden: VS, Verl. für Sozialwiss.
- Brüning, L. & Saum, T. (2009). *Erfolgreich unterrichten durch Kooperatives Lernen*. Band 2. Neue Strategien zur Schüleraktivierung. Essen: NDS-Verlag.
- Bühler, K. (1999). *Sprachtheorie: Die Darstellungsfunktion der Sprache*. 3. Auflage. Stuttgart: Lucius & Lucius.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Second edition. Hillsdale, NJ [u. a.]: Erlbaum.

- Cohen, R.; Rey, R. E. & Sixtil, F. (1969). Die Kombination von Häufigkeit und Intensität im diagnostischen Urteil. In: *Psychologische Forschung*, Vol. 33. S. 9-20.
- Disselkamp, C. (2004). "Sonstige Mitarbeit" - zwischen Angebot und Bringschuld. In: *Der altsprachliche Unterricht Latein, Griechisch*, Bd. 47 H. 6. S. 53-55.
- Dohse, W. (1967). *Das Schulzeugnis: Sein Wesen und seine Problematik* (Vol. 10). Weinheim/Bergstr.: Beltz.
- Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge [u. a.]: Cambridge Univ. Press.
- Fiehler, R. (2009). Mündliche Kommunikation. In: *Becker-Mrotzek, M. (Hrsg.) Mündliche Kommunikation und Gesprächsdidaktik* (Vol. 3). Baltmannsweiler: Schneider-Verl. Hohengehren. S. 25-51.
- Furck, C.-L. (1967). *Das pädagogische Problem der Leistung in der Schule*. 3. unveränderte Auflage. Weinheim: Beltz.
- Göller, A. (1968). *Zensuren und Zeugnisse*. 2. Auflage. Stuttgart: Klett.
- Heckhausen, H. (1974). *Leistung und Chancengleichheit* (Vol. 2). Göttingen: Hogrefe.
- Heckhausen, H. (1980). *Motivation und Handeln: Lehrbuch der Motivationspsychologie*. Berlin [u. a.]: Springer.
- Heller, K. A. (Hrsg.) (1984). *Leistungsdiagnostik in der Schule*. 4., völlig neubearbeitete Auflage. Bern/ Stuttgart/ Toronto: Huber.
- Heller, K. A. & Hany, E. A. (2014). Standardisierte Schulleistungsmessungen. In: *Weinert, F. E. (Hrsg.) Leistungsmessungen in Schulen*. Weinheim [u. a.]: Beltz. S. 87-101.
- Helmke, A. (2014). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. 5. Auflage. Seelze-Velber: Klett/Kallmeyer.
- Helmke, A. & van Aken, M. A. (1995). The causal ordering of academic achievement and selfconcept of ability during elementary school. A longitudinal study. In: *Journal of Educational Psychology*, Bd. 87., S. 624-637.
- Hessisches Kultusministerium (2006). *Verordnungen zum Hessischen Schulgesetz*. http://www.bns.info/images/Schulgesetz_2006.pdf, Zugriff am 17.08.16
- Hessisches Schulgesetz (HSchG) in der Fassung vom 14. Juni 2005 (GVBl. I S. 441), zuletzt geändert durch Gesetz vom 24. März 2015 (GVBl. S. 118). https://kultusministerium.hessen.de/sites/default/files/media/lesefassung_hessisches_schulgesetz_mit_inhaltsverzeichnis_stand_14.04.2016.pdf, Zugriff am 17.08.16
- Holmeier, M. (2013). *Leistungsbeurteilung im Zentralabitur*. Univ., Diss.--Zürich, 2012. Wiesbaden: Springer (Educational Governance, 22).
- Ingenkamp, K. & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik*. 6., neu ausgestattete Auflage. Weinheim [u. a.]: Beltz.
- Jäger, R. S. (2007). *Beobachten, beurteilen, fördern. Lehrbuch für die Aus-, Fort- und Weiterbildung*. Landau: Empirische Pädagogik.

- Jürgens, E. (2010). *Leistung und Beurteilung in der Schule. Eine Einführung in Leistungs- und Bewertungsfragen aus pädagogischer Sicht*. 7., überarb. Aufl. Sankt Augustin: Academia-Verl.
- Jurkowski, S. & Hänze, M. (2016). A closer look at social skills and school performance. Students' peer relations skills and assertion skills as predictors for their written and oral performances. *In: European Journal of Psychology of Education* Jan 18, 2016.
- Kanning, U. P. (2002). Soziale Kompetenz - Definition, Strukturen und Prozesse. *In: Zeitschrift für Psychologie / Journal of Psychology* 210 (4), S. 154–163.
- Kanning, U. P. (2009). *ISK – Inventar Sozialer Kompetenzen*. Göttingen: Hogrefe.
- Kirk, S. (2004). *Beurteilung mündlicher Leistungen. Pädagogische, psychologische, didaktische und schulrechtliche Aspekte der mündlichen Leistungsbeurteilung*. Bad Heilbrunn/Obb.: Klinkhardt.
- Klafki, W. (2007). *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik*. Weinheim [u. a.]: Beltz.
- Klauer, K. J. (2014). Wie misst man Schulleistungen? *In: Weinert, F. E. (Hrsg.). Leistungsmessungen in Schulen*. Weinheim [u. a.]: Beltz. S. 103-115.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2012). Bildungsstandards im Fach Deutsch für die Allgemeine Hochschulreife. http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Deutsch-Abi.pdf, Zugriff am 18.08.16
- Koch, P & Oesterreicher, W. (2007). Schriftlichkeit und kommunikative Distanz. *In: Zeitschrift für germanistische Linguistik*. Band 35 (3). S. 346-375.
- Krapp, A. (1984). Forschungsergebnisse zur Bedingungsstruktur der Schulleistung. *In: Heller, K. A. (Hrsg.) Leistungsdiagnostik in der Schule*. 4., völlig neubearbeitete Auflage. Bern/ Stuttgart/ Toronto: Huber. S. 46-62.
- Krieger, R. (2003). Das sogenannte Mündliche: Eine Befragung von Beurteilten und Beurteilern zur Praxis und Problematik der Bewertung mündlicher Leistungen. *In: Zeitschrift Bildung und Erziehung*. Band 56. S. 75-91.
- Krumwiede, F.; Schneider, J. & Wickner, M.-C. (2014). *Mündliche und praktische Leistungen bewerten. Das Praxisbuch ; Sekundarstufe I + II; Profi-Tipps und Materialien aus der Lehrerfortbildung*. Donauwörth: Auer.
- Langhammer, R. (1997). Mündliche Noten – Ergebnisse einer Umfrage unter Lehrkräften und Schülern. *In: Berthold, S.; Pabst-Weinschenk, M., Wagner, R. W. & Naumann, C. L. (Hrsg.). Sprecherziehung im Unterricht (Vol. 33)*. München [u. a.]: Reinhardt. S. 80-93.
- Langhammer, R. (1998). "Mündliche Noten" im Geschichts- und Politikunterricht. Erhebungs- und Bewertungsvorschläge. *In: Geschichte, Politik und ihre Didaktik, Bd. 26 H. 1-2*. S. 38-50.
- Langhorst, E. (1984). Beobachtung und Beurteilung des Schülerverhaltens im Unterricht. *In: Heller, K. A. (Hrsg.) Leistungsdiagnostik in der Schule*. 4., völlig neubearbeitete Auflage. Bern/ Stuttgart/ Toronto: Huber. S. 208-228.
- Marsh, H. W. & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *In: Journal of Personality and Social Psychology*, 47, S. 213–231.

- Martenstein, H. (2011). "Man erzieht die Leute zu Dauerlerner". Harald Martenstein über eine überschätzte Kulturtechnik. In: *ZEITmagazin Nr. 5/2011*. Online verfügbar unter <http://www.zeit.de/2011/05/Martenstein>, Zugriff am 02.09.2016
- Mietzel, G. (2007). *Pädagogische Psychologie des Lernens und Lehrens*. 8., überarb. und erw. Aufl. Göttingen: Hogrefe (Lehrbuch).
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte. In: *Psychologische Rundschau*, 55 (1), S. 19–27.
- Neuweg, G. H. (2006). *Schulische Leistungsbeurteilung. Rechtliche Grundlagen und pädagogische Hilfestellungen für die Schulpraxis*. 3., aktual. und erw. Aufl., Stand der Rechtslage: 1. September 2006. Linz: Trauner (Ratgeber Schulpraxis).
- Nickel, R. (2004). Beurteilung und Bewertung "mündlicher Mitarbeit". In: *Der altsprachliche Unterricht, Latein, Griechisch. Band 47*. S. 2-7.
- Nipkow, K. E. (1978). Zur Entstehung des Leistungsprinzips in der Geschichte des deutschen Schulwesens. In: *Beckmann, H.-K. (Hrsg.). Leistung in der Schule (Vol. 166)*. Braunschweig: Westermann. S. 7-32.
- Notenbildungsverordnung (NVO) des Kultusministeriums des Landes Baden-Württemberg vom 5. Mai 1983, letzte berücksichtigte Änderung am 19. April 2016. <http://www.landesrecht-bw.de/jportal/?quelle=jlink&query=NotBildV+BW&psml=bsbawueprod.psml&ax=true&aiz=true#jlr-NotBildVBWV6P6-jlr-NotBildVBWV4P6-jlr-NotBildVBWV5P6>, Zugriff am 15.08.2016
- Nuding, A. (2006). *Beurteilen durch Beobachten: Gewinnung diagnostischer Informationen als Grundlage für Beurteilungen* (Vol. 21). Baltmannsweiler: Schneider-Verl. Hohengehren.
- Oberstufen- und Abiturverordnung (OAVO) für Hessen vom 20. Juli 2009 (ABl. S. 408), zuletzt geändert durch Verordnung vom 4. April 2013 (ABl. S. 158). https://verwaltung.hessen.de/irj/servlet/prt/portal/prtroot/slimp.CMReader/HKM_15/HKM_Internet/med/619/6197d14d-9181-f317-9cda-a2b417c0cf46,22222222-2222-2222-2222-222222222222,true, Zugriff am 20.08.16
- OECD (2016). *Low-Performing Students: Why They Fall Behind and How to Help Them Succeed*, PISA, OECD Publishing, Paris. <http://www.oecd-ilibrary.org/docserver/download/9816011e.pdf?expires=1470161397&id=id&ccname=quest&checksum=6AE9A5525EC23A9726506527507489C3>, Zugriff am 02.08.2016
- Paradies, L.; Wester, F. & Greving, J. (2014). *Leistungsmessung und -bewertung*. Berlin: Cornelsen.
- Potthoff, U.; Steck-Lüschow, A. & Zitzke, E. (1996). Mündliche Leistungen bewerten – aber wie? Beobachtung und Bewertung von Gesprächsfähigkeit. In: *Bambach, H. (Hrsg.) Prüfen und beurteilen: Zwischen Fördern und Zensieren (Vol. 14)*. Seelze: Friedrich. S. 14-17.
- Rheinberg, F. (2014). Bezugsnormen und schulische Leistungsbeurteilung. In: *Weinert, F. E. (Hrsg.). Leistungsmessungen in Schulen. Weinheim [u. a.]: Beltz*. S. 59-71.
- Ritz-Fröhlich, G. (1982). *Das Gespräch im Unterricht: Anleitung - Phasen - Verlaufsformen*. Bad Heilbrunn, Obb.: Klinkhardt.

- Rolff, H.-G. (1997). *Sozialisation und Auslese durch die Schule*. Weinheim [u. a.]: Juventa-Verl.
- Rosenthal, R & Jacobson, L. (1966). *Teachers' Expectancies: Determinants Of Pupils' IQ Gains*. In: *Psychological Reports*. Band 19. S. 115–11.
- Sacher, W. (2014). *Leistungen entwickeln, überprüfen und beurteilen*. 6. Auflage. Bad Heilbrunn: Klinkhardt.
- Schröter, G. (1981). *Zensuren? Zensuren! Allgemeine und fachspezifische Probleme*. Baltmannsweiler: Schneider.
- Schulz von Thun, F. (2008). *Miteinander reden: 1. Störungen und Klärungen: Allgemeine Psychologie der Kommunikation*. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl.
- Schwark, W., Weiß, W. W. & Regelein, S. (1987). *Beurteilen und Benoten in der Grundschule : Bestandsaufnahme und Anregungen für die Praxis*. München: Ehrenwirth.
- Souvignier, E. & Förster, N. (2014). Die Bewertung der "sonstigen Mitarbeit": Zum Zusammenhang von Unterricht und Bewertung. In: *Schulmagazin 5 bis 10, Bd. 82 H. 6*. S. 7-14.
- Stevens, S. S. (1966). Mathematics, Measurement and Psychophysics. In: *Stevens, S. S. (Hrsg.) Handbook of experimental psychology*. New York [u. a.]: Wiley. S. 1-49.
- Stolz, W. (2005). *Starthilfe Physik: Ein Leitfaden für Studienanfänger der Naturwissenschaften, des Ingenieurwesens und der Medizin*. Stuttgart [u. a.]: Teubner.
- Tulodziecki, G.; Herzig, B & Grafe, S. (2010). Leistungsmessung und Leistungsbeurteilung. In: *Mägdefrau, J. (Hrsg.) Schulisches Lehren und Lernen : pädagogische Theorie an Praxisbeispielen*. Bad Heilbrunn: Klinkhardt. S. 10-32.
- Valtin, R. (2002). *Was ist ein gutes Zeugnis? Noten und verbale Beurteilungen auf dem Prüfstand*. Weinheim: Juventa.
- Watzlawick, P.; Beavin, J. H. & Jackson, D. D. (2011). *Menschliche Kommunikation. Formen, Störungen, Paradoxien*. 12. unveränd. Aufl. Bern: Huber.
- Weinert, F. (1999a). Aus Fehlern lernen und Fehler vermeiden lernen. In: *Althof, W. (Hrsg.) Fehlerwelten : vom Fehlermachen und Lernen aus Fehlern*. Opladen: Leske + Budrich. *Fehlerwelten. Vom Fehlermachen und Lernen aus Fehlern. Beiträge und Nachträge zu einem interdisziplinären Symposium aus Anlaß des 60. Geburtstags von Fritz Oser*. Opladen: Leske + Budrich. S. 105-116.
- Weinert, F. E. (1999b). Die fünf Irrtümer der Schulreformer. Welche Lehrer, welchen Unterricht braucht das Land? In: *Psychologie heute, 1999 (7)*. S. 28-34.
- Weinert, F. E. (Hrsg.) (2014). *Leistungsmessungen in Schulen*. Weinheim [u. a.]: Beltz.
- Wengert, H. G. (2011). Leistungsbeurteilung in der Schule. In: *Bovet, G. & Huwendiek, V. (Hrsg.) Leitfaden Schulpraxis. Pädagogik und Psychologie für den Lehrerberuf*. 6. Auflage. Berlin: Cornelsen. S. 324-349.

-
- Winter, F. (2016). *Leistungsbewertung. Eine neue Lernkultur braucht einen anderen Umgang mit den Schülerleistungen* (Vol. 49). Baltmannsweiler: Schneider-Verlag Hohengehren GmbH.
- Ziegenspeck, J. (1976). *Zensur und Zeugnis in der Schule*. Hannover: Schroedel.
- Ziegenspeck, J. (1999). *Handbuch Zensur und Zeugnis in der Schule: Historischer Rückblick, allgemeine Problematik, empirische und bildungspolitische Implikationen*. Bad Heilbrunn/Obb.: Klinkhardt.

7. Abbildungsverzeichnis

Abbildung 1: Karikatur aus Souvignier & Förster (2014, S. 7)	1
Abbildung 2: Die vier Seiten (Aspekte) einer Nachricht nach Schulz von Thun (2008, S. 30)	18
Abbildung 3: Kompetenzen der Mündlichkeit nach Abraham (2008, S. 29).....	20
Abbildung 4: Kriterien für die Bewertung mündlicher Mitarbeit aus Krumwiede u. a. (2014, S. 24)	35
Abbildung 5: Verteilung der befragten Lehrkräfte nach Fächern	49
Abbildung 6: Verteilung der befragten Lehrkräfte nach Aufgabenfeldern	49
Abbildung 7: Verteilung der befragten Lehrkräfte nach Berufserfahrung in Jahren	50
Abbildung 8: Benannte Kriterien für die Bewertung mündlicher Mitarbeit	53
Abbildung 9: Rangfolge der Kriterien der gesamten Antworten der Stichprobe	55
Abbildung 10: Rangfolge der Kriterien nach Aufgabenfeldern.....	56
Abbildung 11: Rangfolge der Kriterien nach Berufserfahrung	63
Abbildung 12: Anteil der Lehrkräfte mit Vorgaben an der Schule	67
Abbildung 13: Anteil an Lehrkräften mit Absprachen im Kollegium	68
Abbildung 14: Verteilung der Bezugsnormorientierung der gesamten Stichprobe	69
Abbildung 15: Verteilung der Bezugsnormorientierung nach Berufserfahrung	69
Abbildung 16: Regelmäßigkeit der Dokumentation der gesamten Stichprobe	71
Abbildung 17: Regelmäßigkeit der Dokumentation nach Berufserfahrung	72
Abbildung 18: Genannte Gründe für eine seltenere als gewünschte Dokumentation	73
Abbildung 19: Reflexion der Bewertung mündlicher Mitarbeit	73

8. Tabellenverzeichnis

Tabelle 1: Kategorien für die Beobachtung von Gesprächsfähigkeit nach Potthoff u. a. (1996, S. 14).....	35
Tabelle 2: Kriterien für die Bewertung mündlicher Mitarbeit nach Nickel (2004, S. 3 f.).....	36
Tabelle 3: Abweichung der Relevanz der Qualität der Beiträge nach Aufgabenfeldern	57
Tabelle 4: Abweichung der Relevanz der Kontinuität der Beiträge nach Aufgabenfeldern	58
Tabelle 5: Abweichung der Relevanz der Anzahl der Beiträge nach Aufgabenfeldern	58
Tabelle 6: Abweichung der Relevanz der Gesprächsfähigkeit nach Aufgabenfeldern	59
Tabelle 7: Abweichung der Relevanz der (Lern-) Entwicklung nach Aufgabenfeldern	59
Tabelle 8: Abweichung der Relevanz der Erledigung der Hausaufgaben nach Aufgabenfeldern	60
Tabelle 9: Abweichung der Relevanz des Gesamteindrucks aufgrund schriftlicher Leistungen nach Aufgabenfeldern	60
Tabelle 10: Abweichung der Relevanz der Intro-/ Extraversion nach Aufgabenfeldern	61
Tabelle 11: Abweichung der Relevanz des Sozial-/ Störverhaltens nach Aufgabenfeldern	61
Tabelle 12: Abweichung der Relevanz der Kontinuität der Beiträge nach Berufserfahrung.....	64
Tabelle 13: Abweichung der Relevanz des Gesamteindrucks aufgrund schriftlicher Leistungen nach Berufserfahrung.....	64
Tabelle 14: Abweichung der Relevanz der Intro-/ Extraversion nach Berufserfahrung.....	64

Tabelle 15: Abweichung der Relevanz der Qualität der Beiträge nach Berufserfahrung.....	65
Tabelle 16: Abweichung der Relevanz der Gesprächsfähigkeit nach Berufserfahrung.....	65
Tabelle 17: Abweichung der Relevanz der Berücksichtigung der (Lern-) Entwicklung nach Berufserfahrung	66
Tabelle 18: Abweichung der Relevanz der Anzahl der Beiträge nach Berufserfahrung.....	66
Tabelle 19: Abweichung der Relevanz des Sozial-/ Störverhaltens nach Berufserfahrung.....	66
Tabelle 20: Abweichung der Relevanz der Erledigung der Hausaufgaben nach Berufserfahrung.....	67
Tabelle 21: Abweichung der Gewichtung der individuellen Bezugsnorm nach Berufserfahrung.....	70
Tabelle 22: Abweichung der Gewichtung der sozialen Bezugsnorm nach Berufserfahrung.....	70
Tabelle 23: Abweichung der Gewichtung der kriterialen Bezugsnorm nach Berufserfahrung.....	70

9. Schlusserklärung

Hiermit versichere ich, dass ich die wissenschaftliche Hausarbeit selbstständig verfasst, keine anderen, als die angegebenen Hilfsmittel verwandt und die Stellen, die anderen benutzten Druck- und digitalisierten Werken sowie bildlichen und sonstigen Darstellungen im Wortlaut oder dem Sinn nach entnommen sind, mit Quellenangaben kenntlich gemacht habe.

Fuldabrück, den 03.11.2016 _____

10. Anhang

A. Kriterienkataloge zur Bewertung mündlicher Mitarbeit – Kopiervorlagen.....	1
A.1 Kriterienkatalog für die mündliche Mitarbeit – Gewichtung durch Punkte..	2
A.2 Dialogischer Beurteilungsbogen	3
A.3 Rückmeldebogen für die Bewertung der mündlichen Unterrichtsmitarbeit	4
A.4 Bausteine für die Rückmeldung mündlicher Leistungen	5
B. Online-Fragebogen	6

A. Kriterienkataloge zur Bewertung mündlicher Mitarbeit – Kopiervorlagen

Die folgenden Kopiervorlagen stammen aus:

Krumwiede, F.; Schneider, J. & Wickner, M.-C. (2014). Mündliche und praktische Leistungen bewerten. Das Praxisbuch ; Sekundarstufe I + II; ProfiTipps und Materialien aus der Lehrerfortbildung. Donauwörth: Auer.

A.1 Kriterienkatalog für die mündliche Mitarbeit – Gewichtung durch Punkte

Name: Datum:				Bemerkungen	
<i>immer wichtig</i>	Die Beiträge sind fachlich richtig.	ja	2		
		teils teils	1		
		nein	0		
	Die mündliche Mitarbeit erfolgt regelmäßig und selbstständig.	meistens	1		
		nein	0		
	Die vereinbarten Gesprächsregeln werden eingehalten.	ja	1		
		eher nein	0		
	Die Mitarbeit ist verständlich und die Beiträge sind nachvollziehbar.	ja	2		
		teils teils	1		
		nein	0		
	Die Antwort erfolgt in vollständigen Sätzen.	meistens	1		
		selten	0		
<i>diesmal wichtig</i>	Das Interesse am Thema wird deutlich.	sehr häufig	2		
		manchmal	1		
		selten	0		
	Die Begründung der eigenen Meinung findet	ausführlich statt.	3		
		genau statt.	2		
		statt.	1		
		nicht statt.	0		
Du hast von möglichen Punkten Punkte erreicht.					
Note:					

A.2 Dialogischer Beurteilungsbogen

Name:		Ich		Mein/e Lehrer/in	
		😊	☹️	😊	☹️
immer wichtig	Ich zeige auf, wenn ich etwas weiß.				
	Ich kann abwarten, bis ich aufgerufen werde.				
	Ich kann anderen den Vortritt lassen.				
	Ich melde mich regelmäßig.				
	Ich kann etwas sagen, das zum Thema passt.				
	Ich kann Gesprächsregeln beachten.				
diesmal wichtig	Ich kann anderen etwas erklären.				
	Ich kann meine eigene Meinung begründen.				
	Ich kann in ganzen Sätzen antworten.				
	Ich kann gezielt Fragen stellen.				
	Ich kann mein Vorwissen mit einbringen.				
	Ich kann Gelerntes wiedergeben.				
	Notenvorschlag				
Bemerkungen:					

A.3 Rückmeldebogen für die Bewertung der mündlichen Unterrichtsmitarbeit

+ trifft in vollem Maße zu o trifft teilweise zu – trifft eher nicht bis gar nicht zu

Mündliche Mitarbeit im Unterricht	+	o	-
So meldest du dich:			
Du rufst nicht dazwischen.			
Du kannst abwarten.			
Du kannst anderen den Vortritt lassen.			
Du meldest dich regelmäßig.			
So nimmst du Rücksicht auf die Klasse:			
Du beachtest Gesprächsregeln.			
Du kannst anderen etwas erklären.			
Du respektierst die Meinungen anderer.			
So sind deine Beiträge:			
Deine Beiträge passen zum Thema.			
Du kannst gezielt Fragen stellen.			
Du kannst dein Vorwissen einbringen.			
Du kannst Inhalte wiedergeben.			
Du kannst das Thema weiterführen.			
Du kannst deine Meinung begründen.			
Wenn du unaufgefordert aufgerufen wirst, ...			
zeigst du, dass du aufmerksam zugehört hast.			
kannst du richtige Antworten geben.			
So kannst du zuhören:			
Du kannst aufmerksam und konzentriert zuhören.			
Du gehst auf das ein, was andere vor dir gesagt haben.			
Du gibst Gehörtes vollständig und richtig wieder.			
Du artikulierst deine mündliche Mitarbeit:			
ideenreich			
wortgewandt			
Insgesamt			
Note			

A.4 Bausteine für die Rückmeldung mündlicher Leistungen

<p>Die Schülerin oder der Schüler arbeitet nicht freiwillig im Unterricht mit. Nach Aufforderung sind die Äußerungen falsch.</p>	Note: 6
<p>Die Schülerin oder der Schüler arbeitet nicht freiwillig im Unterricht mit. Nach Aufforderungen sind die Äußerungen nur selten und nur in Teilen richtig.</p>	Note: 5
<p>Die Schülerin oder der Schüler arbeitet nur gelegentlich freiwillig im Unterricht mit. Die Wortmeldungen erfolgen in Form einer Reproduktion von Fakten, die mit dem derzeitigen Unterrichtsthema zusammenhängen. Die Äußerungen sind im Wesentlichen richtig.</p>	Note: 4
<p>Die Schülerin oder der Schüler arbeitet regelmäßig freiwillig im Unterricht mit. Die Wortmeldungen erfolgen in Form einer Reproduktion von Fakten, die mit dem derzeitigen Unterrichtsthema zusammenhängen. Die Schülerin oder der Schüler schafft es, Äußerungen mit Kenntnissen des Stoffes der gesamten Unterrichtsreihe zu verknüpfen.</p>	Note: 3
<p>Die Schülerin oder der Schüler arbeitet regelmäßig freiwillig im Unterricht mit. Die Wortmeldungen erfolgen in Form einer Reproduktion von Fakten, die mit dem derzeitigen Unterrichtsthema zusammenhängen. Die Schülerin oder der Schüler schafft es, Äußerungen mit Kenntnissen des Stoffes der gesamten Unterrichtsreihe zu verknüpfen. Das Verständnis schwieriger Sachverhalte und deren Einordnung in den Gesamtzusammenhang des Themas werden deutlich. Problemlösungskompetenz ist vorhanden. Die Schülerin oder der Schüler ist in der Lage, zwischen Wesentlichem und Unwesentlichem zu unterscheiden. Es sind Kenntnisse vorhanden, die über die Unterrichtsreihe hinausreichen.</p>	Note: 2
<p>Die Schülerin oder der Schüler arbeitet regelmäßig freiwillig im Unterricht mit. Die Wortmeldungen erfolgen in Form einer Reproduktion von Fakten, die mit dem derzeitigen Unterrichtsthema zusammenhängen. Die Schülerin oder der Schüler schafft es, Äußerungen mit Kenntnissen des Stoffes der gesamten Unterrichtsreihe zu verknüpfen. Probleme werden erkannt und in einen größeren Zusammenhang gebracht. Die Schülerin oder der Schüler ist in der Lage, einen Sachverhalt sachgerecht und fundiert zu beurteilen. Eine eigenständige gedankliche Leistung wird als Beitrag zur Problemlösung deutlich. Die Darstellung der mündlichen Mitarbeit erfolgt in angemessener, klarer sprachlicher Form.</p>	Note: 1

B. Online-Fragebogen

Die Bewertung mündlicher Mitarbeit im Unterricht

Seite 1

Im Rahmen meiner Examensarbeit an der Universität Kassel möchte ich mit Ihrer Hilfe herausfinden, wie die Bewertung der mündlichen Mitarbeit an ausgewählten Kasseler Gymnasien stattfindet. Dazu bitte ich Sie, den folgenden Fragebogen auszufüllen.

Die Bearbeitung des Fragebogens kann 5 -10 Minuten in Anspruch nehmen.

Die Daten dienen ausschließlich zur anonymen Auswertung im Rahmen der Examensarbeit und werden NICHT an Dritte weitergegeben!

Seite 2

1. Persönliche Angaben

Geschlecht *

männlich

weiblich

An welcher Schule sind Sie tätig? *

Welche Fächer unterrichten Sie? *

Bitte zutreffende Felder anklicken. Sie können bis zu drei Fächer angeben.

Mathematik

Deutsch

Englisch

Französisch

Geschichte

Politik und Wirtschaft

Geografie

- Kunst
- Musik
- Physik
- Chemie
- Biologie
- Informatik
- Spanisch
- Italienisch
- Sport
- evangelische Religion
- katholische Religion
- Ethik
-

Wieviele Berufsjahre haben Sie bereits vollendet? *

Seite 3

2. Schule und Kollegium

Gibt es an Ihrer Schule Vorgaben zur Bewertung mündlicher Mitarbeit? *

- ja
- nein

Wenn ja, welche?

Seite 5

3. Kriterien der Bewertung mündlicher Mitarbeit

Bitte wählen Sie eines Ihrer Fächer aus und beantworten, in Bezug darauf, die untenstehende Frage. *

- Mathematik
- Deutsch
- Englisch
- Französisch
- Geschichte
- Politik und Wirtschaft
- Geografie
- Kunst
- Musik
- Physik
- Chemie
- Biologie
- Informatik
- Spanisch
- Italienisch
- Sport
- evangelische Religion
- katholische Religion
- Ethik
-

Für wie wichtig erachten Sie folgende Kriterien zur Beurteilung mündlicher Mitarbeit? *

Bitte nehmen Sie durch das Zuweisen von Ziffern bzw. das Verschieben (dazu die Pfeile anklicken) der Begriffe eine Hierarchisierung vor. Vom wichtigsten Kriterium absteigend.

- ⬆ Berücksichtigung der Intro-/ Extraversion (Temperament)
- ⬆ Berücksichtigung des Sozial-/ Störverhaltens
- ⬆ Gesamteindruck der Leistung (aufgrund von z. B. schriftlichen Leistungen)
- ⬆ Kontinuität der Beiträge (Frequenz)
- ⬆ Berücksichtigung der (Lern-)Entwicklung
- ⬆ Anzahl der Beiträge
- ⬆ Gesprächsfähigkeit (verbale Ausdrucksfähigkeit)
- ⬆ Fachliche Qualität der Beiträge
- ⬆ Erledigung von Hausaufgaben

Seite 6

3. Kriterien der Bewertung mündlicher Mitarbeit

Möchten Sie für Ihr zweites Fach eine andere Hierarchisierung vornehmen? *

JA --> Die Hierarchisierung der Kriterien ist in meinen Fächern unterschiedlich.

NEIN --> Die Hierarchisierung der Kriterien ist in meinen Fächern gleich.

- ja (Weiterleitung auf Seite 7)
- nein (Weiterleitung auf Seite 8)

Seite 7

3. Kriterien der Bewertung mündlicher Mitarbeit

Bitte wählen Sie Ihr anderes Fach aus und beantworten, in Bezug darauf, die untenstehende Frage. *

- Mathematik
- Deutsch
- Englisch
- Französisch

-
- Geschichte
 - Politik und Wirtschaft
 - Geografie
 - Kunst
 - Musik
 - Physik
 - Chemie
 - Biologie
 - Informatik
 - Spanisch
 - Italienisch
 - Sport
 - evangelische Religion
 - katholische Religion
 - Ethik
 -

Für wie wichtig erachten Sie folgende Kriterien zur Beurteilung mündlicher Mitarbeit? *

Bitte nehmen Sie durch das Zuweisen von Ziffern bzw. das Verschieben (dazu die Pfeile anklicken) der Begriffe eine Hierarchisierung vor. Vom wichtigsten Kriterium absteigend.

- Fachliche Qualität der Beiträge
- Berücksichtigung des Sozial-/ Störverhaltens
- Gesprächsfähigkeit (verbale Ausdrucksfähigkeit)
- Anzahl der Beiträge
- Kontinuität der Beiträge (Frequenz)
- Berücksichtigung der (Lern-)Entwicklung
- Berücksichtigung der Intro-/ Extraversion (Temperament)
- Erledigung von Hausaufgaben
- Gesamteindruck der Leistung (aufgrund von z .B. schriftlichen Leistungen)

Seite 8

4. Dokumentation der mündlichen Mitarbeit

Wie häufig halten Sie die mündlichen Leistungen der Lernenden schriftlich fest? *

- nach jeder Stunde
- täglich
- wöchentlich
- monatlich
- halbjährlich

Würden Sie gerne häufiger die Leistungen Ihrer Lernenden dokumentieren? *

- ja (Weiterleitung auf Seite 9)
- nein (Weiterleitung auf Seite 10)

Seite 9

4. Dokumentation der mündlichen Mitarbeit

Beschreiben Sie kurz, wie die Diskrepanz zustande kommt. *

Dies kann auch stichwortartig geschehen.

Seite 10

4. Dokumentation der mündlichen Mitarbeit

Halten Sie die mündlichen Leistungen der SchülerInnen in einer Tabelle fest? *

ja (Weiterleitung auf Seite 11)

nein (Weiterleitung auf Seite 12)

Seite 11

4. Dokumentation der mündlichen Mitarbeit

Sie benutzen also eine Tabelle, um die Leistungen festzuhalten. Was tragen Sie dort ein? *

++/+/0/--

Noten (-punkte)

Seite 12

4. Dokumentation der mündlichen Mitarbeit

Optional: Hier ist Platz für Anmerkungen bezüglich Ihres persönlichen Vorgehens bei der Dokumentation.

Seite 13

5. Persönliche Reflexion der Bewertungspraxis

Inwiefern treffen diese Aussagen auf Sie zu? *

trifft nicht zu trifft eher nicht zu teils / teils trifft eher zu trifft völlig zu

Ich bin davon
überzeugt, dass
eine andere
Lehrperson im
selben Unterricht zu
ähnlichen
Bewertungen
kommen würde.

Wenn ich den
Beobachtungs-
prozess
wiederholen könnte,
würde ich zu
ähnlichen
Ergebnissen
kommen.

Meine Bewertung
der mündlichen
Mitarbeit beschreibt
die Leistungen der
Lernenden
zutreffend.

Die Kriterien meiner
Bewertung stelle ich
den Lernenden im
Vorfeld transparent
dar.

Die Lernenden sind
mit meiner
Bewertung der
mündlichen
Mitarbeit
einverstanden.

Ich bewerte die mündliche Mitarbeit des Lernenden anhand... *

	trifft nicht zu	trifft eher nicht zu	trifft eher zu	trifft völlig zu
... meines vorher festgelegten Erwartungs- horizonts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... des Vergleichs mit der Leistung anderer in der Lerngruppe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... der vorherigen Leistungen des Lernenden (Lernfortschritt).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Seite 14

Fast geschafft

Zum Schluss: Anmerkungen & Fragen

Haben Sie Anmerkungen / Fragen zum Thema "Bewertung mündlicher Mitarbeit"?

Gibt es Schwierigkeiten bei der Benotung mündlicher Mitarbeit?

Haben Sie Interesse an den Ergebnissen dieser Umfrage?

Eine kurze Zusammenfassung der anonymen Auswertung stelle ich gerne zur Verfügung.

Wenn Interesse besteht, bitte ich Sie, hier Ihren Namen einzutragen ODER mir eine kurze E-Mail zu senden.

Seite 15

Fertig

WICHTIG: Klicken Sie auf "Fertig", um die Umfrage abzuschließen.

Die Umfrage ist nun beendet. Vielen Dank für Ihre Teilnahme.

» Umleitung auf Schlussseite von Umfrage Online ([ändern](#))