



Beurteilungen von Förderanträgen durch individuelle Gutachtende und durch Gutachtergruppen

Dr. Torsten Futterer



Studien zum Qualitätsmanagement im Hochschulwesen, Nr. 6

Bibliografische Informationen der Deutschen Nationalbibliothek

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

Studien zum Qualitätsmanagement im Hochschulwesen, Nr. 6

ISSN 2750-7475



Dieses Werk ist open access unter der Creative-Commons-Lizenz CC BY-NC-SA verfügbar.

© Zentrale Evaluations- und Akkreditierungsagentur Hannover (ZEVA), Hannover 2024

www.zeva.org

Geschäftsführung: Henning Schäfer (schaefer@zeva.org)

Wissenschaftliche Leitung (kommissarisch): Prof. Dr. Katja Koch & Dr. Marion Rieken



Inhalt

Zusammenfassung.....	2
1 Gegenstand und Ziel dieses Berichts.....	3
2 Das Begutachtungsverfahren zur Förderlinie Innovation Plus.....	3
3 Die Begutachtungsmethode.....	4
4 Individuelle Gutachterurteile und Gruppenurteile im mehrstufigen Verfahren.....	7
5 Individuelle Beurteilungskriterien (Gutachterbias).....	9
6 Bewertungsanpassung durch die Gutachtersitzung.....	11
7 Fazit.....	14
Anhang: Online-Beurteilungsbogen.....	17

Zusammenfassung

Die ZEVA führte in den Jahren 2018 bis 2022 im Auftrag des Niedersächsischen Ministeriums für Wissenschaft und Kultur ein gutachterbasiertes Auswahlverfahren von Förderanträgen niedersächsischer Hochschulen durch. Dafür wurde ein zweistufiges Verfahren gewählt: Aufbauend auf individuelle Gutachterbewertung über einen Online-Fragebogen wurde eine finale Beurteilung durch eine große Gutachtergruppe und mehrere kleine Gutachterteams vorgenommen. Durch dieses gestufte Verfahren, das mit intensiven Abstimmungsprozessen in der Großgruppe und in den Teams verbunden war, konnte ein qualitativ hochwertiger, aber auch aufwändiger Auswahlprozess etabliert werden. Die Begutachtungen waren verbunden mit iterativen Prozessen von individuellen Urteilen zu Gruppenurteilen und führten insgesamt zu einer hohen Zufriedenheit bei den Gutachterinnen und Gutachtern.



1 Gegenstand und Ziel dieses Berichts

Ziel dieses Berichts ist es, ein mehrstufiges Begutachtungsverfahren zur Beurteilung von Förderanträgen vorzustellen und insbesondere die Beurteilung durch Gutachtende auf individueller Ebene und der nachfolgenden Einbettung in Gutachterteams zu beleuchten. Es handelt sich dabei nicht um ein Forschungsdesign zur Untersuchung der Urteilsbildung von Gutachtern und Gutachtergruppen, sondern ausschließlich um Beobachtungen und eine nachträgliche Analyse von Daten, die im Rahmen eines strukturierten Auswahlprozesses von Förderanträgen angefallen sind. Es kann den hohen Anspruch an Forschungsdaten somit nicht erfüllen. Wegen des Fehlens eines Forschungsdesigns und der damit verbundenen eingeschränkten Aussagefähigkeit der Ergebnisse wurde auch auf die Verwendung statistischer Tests verzichtet. Allgemeingültige Aussagen zum Gutachterverhalten werden also nicht abgeleitet. Stattdessen soll ein Einblick in das Vorgehen bei der Begutachtung und das konkrete Verhalten der Gutachtenden gegeben werden.

2 Das Begutachtungsverfahren zur Förderlinie Innovation Plus

Das Land Niedersachsen hatte für die Studienjahre 2018/19 bis 2023/24 Projektmittel zur Entwicklung von innovativen Lehr- und Lernkonzepten ausgeschrieben. Das Förderprogramm „Innovation Plus“ wurde in fünf separaten Förderrunden für fünf aufeinanderfolgende Jahre ausgeschrieben. Die ZEVA wurde beauftragt, von den Förderanträgen der Hochschulen diejenigen zu identifizieren, die sich für eine Projektförderung in besonderem Maße eignen. Die Beurteilung sollte über die Expertise von Gutachterinnen und Gutachtern erfolgen.

Die Projektförderungen wurde jeweils vom Niedersächsischen Ministerium für Wissenschaft und Kultur (MWK Niedersachsen) bei den staatlichen Hochschulen in Niedersachsen (13 Universitäten und 7 Fachhochschulen) ausgeschrieben. In jeder Ausschreibung wurden die inhaltlichen Ziele für die Förderung und die Förderbedingungen konkret ausgeführt. In jeder Förderrunde stand ein Gesamtbetrag von 3 Mio. EUR zur Verfügung, jedes einzelne Projekt konnte mit bis zu 50.000 EUR gefördert werden. Daraus ergab sich je nach Förderrunde eine Anzahl von 62 bis 73 zu fördernden Projekten, die aus 139 bis 155 Anträgen auszuwählen waren. Die ZEVA hatte für jede Runde eine Förderempfehlung abgegeben, die Entscheidung zur Projektförderung verblieb jedoch beim MWK Niedersachsen. Das MWK Niedersachsen ist der Empfehlung in jeder Förderrunde gefolgt.

Die Begutachtungsmethode wird nachfolgend im Detail beschrieben.



3 Die Begutachtungsmethode

Für die Auswahl von geeigneten Projektanträgen wurde ein **gutachterbasiertes Verfahren** gewählt. Um verschiedene Perspektiven auf die Projekte gewährleisten zu können, wurden folgende Personengruppen in die Gutachtergruppe aufgenommen:

- 7 professorale Lehrende mit Erfahrung in der Gestaltung von Studium und Lehre
- 7 Personen aus hochschuldidaktischen Einrichtungen
- 7 Studierende mit Erfahrung in der Gestaltung von Studium und Lehre

Die Personen wurden in **Teams** zusammengeführt, mit jeweils einer Person aus jeder Gruppe. Es ergaben sich also 7 Dreier-Teams mit je einer professoralen Lehrkraft, einer Person aus der Hochschuldidaktik und einem Studenten oder einer Studentin (siehe Abbildung 1). In jeder der Förderrunden wurden die Teams personell neu zusammengestellt, so dass daraus neue Konstellationen entstanden. Im Laufe der aufeinanderfolgenden Förder- und Begutachtungsrunden war ein großer Teil der gesamten Gutachtergruppe mit einer stabilen personellen Besetzung angetreten, nur einzelne Positionen mussten im Laufe des Verfahrens neu besetzt werden.

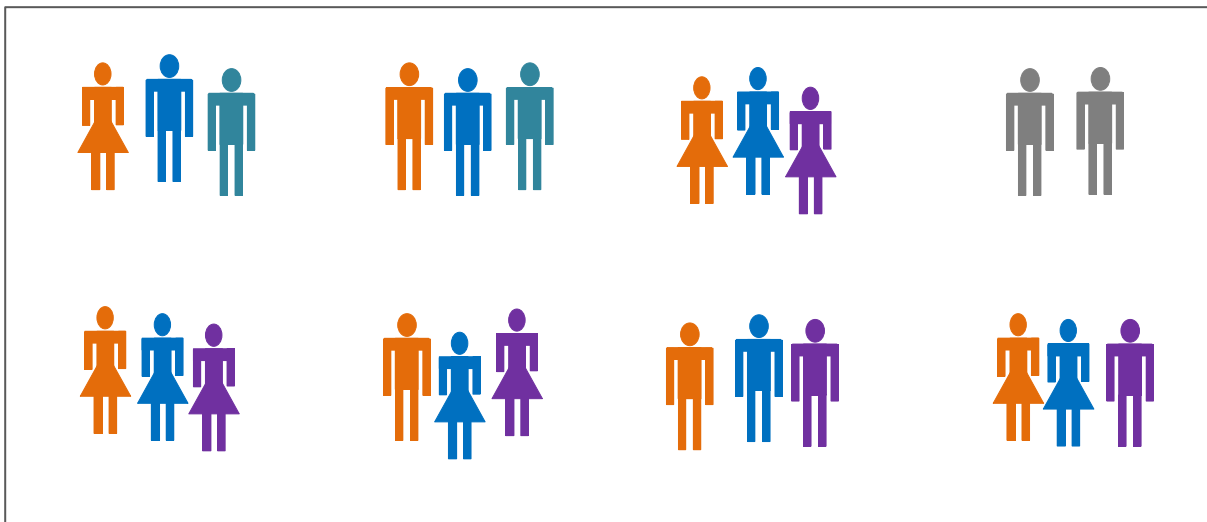


Abbildung 1: Gutachtergruppe von 21 Personen, aufgeteilt auf 7 Teams mit je 3 Personen. In jedem Team waren professorale Lehrende (orange), die Hochschuldidaktik (blau) und Studierende (violett) vertreten. 2 ZEvA-Referenten (grau) betreuten das Verfahren organisatorisch. (Die Darstellung der Geschlechter entspricht der Zusammensetzung der Gutachtergruppe und der Teams in der Begutachtung zur Förderrunde 1.)

Zum Auftakt der jeweiligen Begutachtungsrunde wurden zunächst die Dreier-Teams zusammengestellt, diesen Teams wurden dann (je nach Runde) 17-22 Förderanträge als PDF-Dateien übermittelt. Jede Begutachtungsrunde wurde in zwei aufeinanderfolgende Phasen unterteilt:

- eine individuelle Beurteilung der Gutachtenden über einen Online-Bewertungsbogen *und*
- eine Gutachtersitzung mit Abstimmung der Bewertungen in der Großgruppe und in den Teams.

Die individuelle **Online-Bewertung** erfolgte über einen Beurteilungsbogen, der mit dem Befragungssystem SurveyMonkey generiert wurde. Zur Bearbeitung der Anträge hatten die Gutachtenden ca. vier



Wochen Zeit. Die Ergebnisse der Online-Bewertung wurden von den ZEVA-Referenten analysiert und an die Gutachterteams übermittelt.

Etwa eine Woche später kamen die Gutachtenden in den Runden 1 und 2 zu einer zweitägigen **Gutachtersitzung** (zwei halbe Tage) in Hannover zusammen. In den Runden 3-5 wurden die Sitzungen als Videokonferenzen durchgeführt. In Abbildung 2 ist der Ablauf einer Gutachtersitzung beispielhaft dargestellt.

Montag, 18. November 2019	
bis 14.00 Uhr	Anreise
14.00-15.00 Uhr	Begrüßung Vorstellungsrunde Darstellung der Ergebnisse der Online-Bewertung Verständigung über die Beurteilungskriterien
15.00-18.00 Uhr	Arbeit in Kleingruppen (7 Beurteilungsteams mit je 3 Personen) Diskussion und Abgleich von Beurteilungen auf Basis der vorliegenden Online-Bewertungen
ab 19.00 Uhr	gemeinsames Abendessen
Dienstag, 19. November 2019	
09.00-12.00 Uhr	In der Gesamtgruppe Finale Festlegung der Beurteilung für alle Anträge Mit Fokussierung auf Anträge... <ul style="list-style-type: none">• mit sehr unterschiedlichen Einzelbewertungen• für die sich auch in der Kleingruppendiskussion keine übereinstimmende Bewertung ergab
ab 12.00 Uhr	Möglichkeit der Teilnahme an einem gemeinsamen Mittagessen * anschließend Abreise

Abbildung 2: Ablauf einer Gutachtersitzung

Der **Online-Bewertungsbogen** wird im Anhang zu diesem Bericht wiedergegeben. Er ist wie folgt strukturiert: Nach einigen Hinweisen zur Nutzung des Beurteilungsbogens waren Angaben zur gutachtenden Person und zum begutachteten Projekt zu erfassen. Anschließend wurde die Beurteilung vorgenommen, die sich an der Ausschreibung des Ministeriums orientierte. Die zu berücksichtigenden Themenfelder und die in der Ausschreibung genannten Auswahlkriterien wurden in skalenbasierte Beurteilungssitems umgesetzt, so dass die Gutachtenden systematisch durch die Förderanträge geleitet wurden. Zusätzlich sollte die Nachhaltigkeit der geplanten Maßnahmen eingeschätzt werden.



Entscheidend für das Ergebnis war jedoch die *Gesamtbeurteilung* als *Fazit* der Online-Bewertung. Hier war eine zusammenfassende Beurteilung auf einer vierstufigen Skala abzugeben. Jeder Stufe wurde zudem ein Punktwert zugewiesen:

Das Projekt ...

- ... sollte nicht gefördert werden (= 0 Punkte)
- ... das Projekt ist nur eingeschränkt zur Förderung geeignet (= 1 Punkt)
- ... sollte gefördert werden (= 2 Punkte)
- ... sollte vorrangig gefördert werden (= 3 Punkte)

Die jeweiligen Beurteilungen wurden zudem in einem Freitextfeld kurz begründet.

Für jedes Projekt wurde die **Punktsumme** aus den drei Beurteilungen eines Teams gebildet, wodurch sich ein Wert von 0 bis 9 Punkten ergeben konnte. Der Gutachtergruppe wurde nach Abschluss der Online-Beurteilung eine Übersicht über alle Projektanträge mit allen Gesamtbeurteilungen (inkl. Begründungstext) übermittelt. Dadurch war eine intensive Vorbereitung auf die nachfolgende Gutachtersitzung möglich.

Die **Gutachtersitzungen** waren in vier Abschnitte unterteilt:

- (1) Zunächst hat die ZEvA die Ergebnisse der Online-Bewertung überblicksartig vorgestellt. Berichtet wurde auch, wie viele Projektanträge bereits durch die Online-Bewertung als erfolgreich identifiziert werden konnten (inkl. Ausschöpfung der Gesamtfördersumme).
- (2) Im Anschluss wurden in der Großgruppe Fragen zum weiteren Vorgehen geklärt und insbesondere die Beurteilungskriterien und deren Anwendung auf konkrete Fälle besprochen.
- (3) Die Teams hatten sich daraufhin zur Kleingruppenberatung zurückgezogen, um die ihnen zugewiesenen Projektanträge und die Bewertungen zu besprechen. Es wurde vereinbart, den Fokus auf *die* Projekte zu legen, die große Differenzen bei den individuellen Beurteilungen aufwiesen (z.B.: 0 Punkte – 1 Punkt – 3 Punkte) oder die sich auf der Grenze zur Förderungswürdigkeit befanden (*Gesamtpunktzahl = 5 oder 6*). Projekte mit sehr hohen (7-9 Punkte) oder sehr niedrigen Bewertungen (0-3 Punkte) wurden nur in Einzelfällen erneut betrachtet, da die Übereinstimmung in den Bewertungen dort sehr hoch war. Die Bewertungen konnten von den Gutachtern im Rahmen dieser Kleingruppenabstimmung nun noch einmal angepasst werden. Die Teams erstellten eine Übersicht über die ihnen zugewiesenen Projekte, aus der sowohl die finalen Punktwerte als auch Feedbacktexte für die Antragsteller ersichtlich waren. Die Feedbacktexte waren besonders wertvoll für abgelehnte Anträge, so dass sich die Antragsteller ggf. in der nächsten Förderrunde mit überarbeiteten Anträgen erneut bewerben konnten.
- (4) Die Ergebnisse der Kleingruppenberatungen wurden zusammengetragen, und es wurde festgestellt, ob die Gesamtfördersumme annähernd erreicht wurde. Eine Übersicht über die Förderempfehlungen wurde erstellt und anschließend an das MWK Niedersachsen übermittelt. Die Förderentscheidung wurde den antragstellenden Hochschulen direkt vom MWK mitgeteilt, inklusive eines Feedbacktextes zu den einzelnen Projektanträgen.

Am Ende der jeweiligen Begutachtungssitzung wurden von der Gutachtergruppe Empfehlungen an das MWK Niedersachsen formuliert, die geeignet erschienen, die nachfolgenden Förderrunden zu optimieren. Es wurde empfohlen, diese in die jeweils nächste Förderausschreibung einfließen zu lassen. (Beispiele dafür sind: eine stärkere Fokussierung auf innovative Prüfungsformen und verbindliche



Vorgaben bei der Gestaltung der Förderanträge.) Diese Rückmeldung der Gutachtergruppe wurde dem MWK Niedersachsen gemeinsam mit den Förderempfehlungen übermittelt.

4 Individuelle Gutachterurteile und Gruppenurteile im mehrstufigen Verfahren

In diesem Abschnitt soll erörtert werden, wie sich das **mehrstufige Begutachtungsverfahren** auf die Beurteilung der Förderanträge ausgewirkt hat. Das Verfahren, das in diesem Fall explizit angewendet wurde, liegt implizit auch in den Peer-Review-Verfahren zur Akkreditierung von Studiengängen oder zur Systemakkreditierung von Hochschulen vor. Die dort tätigen Gutachtergruppen nehmen in der Regel eine individuelle Vorab-Bewertung anhand des Selbstberichts der Hochschule vor („Einschätzung des Selbstberichts“). Diese wird jedoch oft nicht systematisch erfasst, sondern fließt nur in die Gruppenbeurteilung bei den Vor-Ort-Gesprächen an der Hochschule ein. Bei dem hier explizit angewendeten mehrstufigen Verfahren ergab sich die Gelegenheit zu einer intensiveren Analyse der Methode, die sich bei impliziter Anwendung nicht bietet.

Ein Vergleich der gutachterlichen Bewertungen bietet sich immer dann an, wenn innerhalb eines Teams identische Anträge von verschiedenen Personen unabhängig voneinander geprüft werden. Dies war bei der Online-Bewertung für die Innovation-Plus-Anträge der Fall: Jeder Antrag wurde von drei Gutachtenden eines Teams unabhängig bewertet. Ideal wäre es für die Analyse gewesen, wenn *alle* Anträge von *allen* Gutachtenden bewertet worden wären. Angesichts der hohen Anzahl von Anträgen war dies aber nicht leistbar – die Verfahrensökonomie war in diesem Fall vorrangig. Die Bewertungen der Gutachtenden und der Teams wurden ausschließlich quantitativ analysiert, und zwar über die Punktwerte, die für den jeweiligen Förderantrag vergeben wurden. Darüber hinaus wurde die Punktschritte betrachtet, die die einzelnen Gutachtenden für die ihnen zugewiesenen Anträge vergeben haben.

Die **Entwicklung der Beurteilung** über die zwei Stufen des Verfahrens, also von der Online-Bewertung bis zum Ende der Gutachtersitzung ist den Abbildungen 3 und 4 zu entnehmen. In Abbildung 3 ist exemplarisch das Ergebnis von Runde 4 dargestellt. Als förderungswürdig wurden die Anträge mit einer Punktschritte von 6 und mehr Punkten angesehen. Diese befinden sich links von der vertikalen rot gestrichelten Linie, die beschriftet ist mit „60 Anträge = 2,85 Mio. €“. Dies beziffert die Anzahl zur Förderung empfohlener Anträge und die erreichte Gesamtfördersumme.

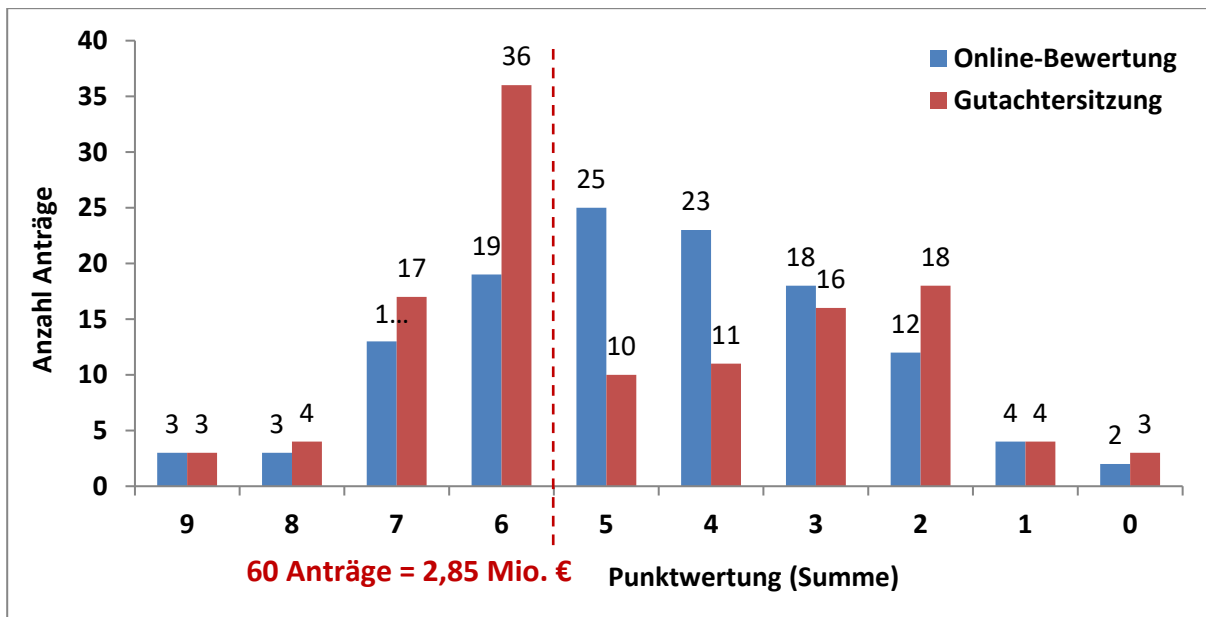


Abbildung 3: Bewertung der 4. Förderrunde

Während die Online-Bewertung (blaue Balken) in etwa einer Normalverteilung folgt, ist die Verteilung nach der Gutachtersitzung zweigipflig. Es ist somit gelungen, einen Kontrast zwischen den akzeptierten Anträgen (6-9 Punkte) und den abgelehnten Anträgen (0-5 Punkte) zu erzeugen. Das gestufte Merkmal der „Punktsumme“ wurde in ein dichotomes Merkmal der „Förderungswürdigkeit“ überführt. Deutlich zu erkennen ist eine Verschiebung bei der Anzahl der Anträge, die mit 4, 5 oder 6 Punkten bewertet wurden. Neubewertungen bei sehr hohen oder niedrigen Punktwerten gab es nur dann, wenn die individuellen Gutachterwertungen sehr unterschiedlich ausfielen. Es fällt auf, dass die Bewertungen nach der Gutachtersitzung insgesamt positiver ausfielen als vorher; ein Trend, der allerdings durch die selbstgesteckte „Fördergrenze“ von 6 Punkten forciert wurde.

In Abbildung 4 werden die Ergebnisse über alle fünf Bewertungsrounds in Summe dargestellt. Die für Abbildung 3 beschriebenen Trends sind auch hier zu erkennen.

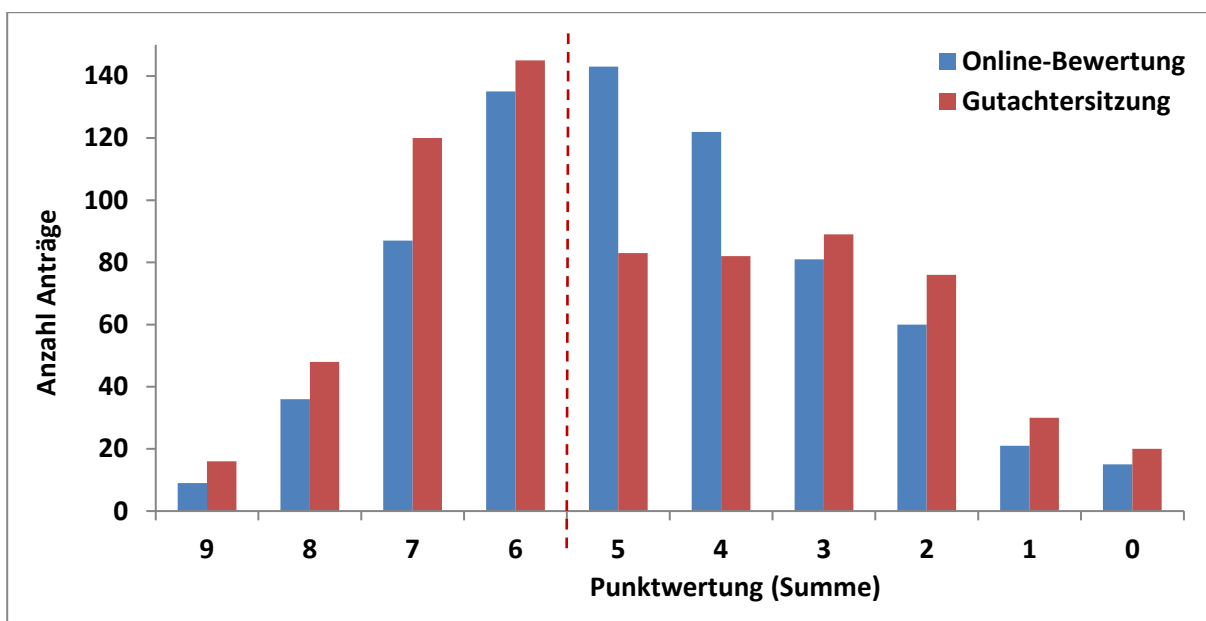


Abbildung 4: Summe der Bewertungen der Förderrunden 1-5



5 Individuelle Beurteilungskriterien (Gutachterbias)

Die Daten der **Online-Bewertung** ermöglichen es, Aussagen zum Gutachterbias zu machen. In den Daten zeigt sich eine Tendenz der Gutachtenden zu eher positiven (milderen) oder eher negativen (schärferen) Urteilen.

In allen nachfolgenden Diagrammen werden (beispielhaft) ausschließlich Daten aus der **ersten Förderrunde** herangezogen, für die weiteren Förderrunden zeigen sich jedoch ähnliche Muster. Außerdem handelt es sich durchweg um die unabhängigen Erstbewertungen im **Online-Verfahren**.

In Abbildung 5 ist zunächst die Summe der Punkte dargestellt, die von den 7 Teams für die jeweils 22 zu beurteilende Förderanträge vergeben wurden. Diese lagen zwischen 82 und 119 Punkten bei identischer Anzahl von 22 Anträgen je Team. Die Aussagekraft der Daten ist allerdings begrenzt, da die Teams mit unterschiedlichen Anträgen versorgt wurden und nicht sichergestellt war, dass die 22 Anträge, die den Teams jeweils zugeordnet wurden, im Mittel die gleiche Qualität oder Güte (Förderungswürdigkeit) aufwiesen.

Abbildung 6 zeigt die Punktschichten für die einzelnen Gutachtenden: Sie erreichen Punktschichten von 22 bis 45 Punkten. Aber auch hier waren die Anträge nicht identisch.

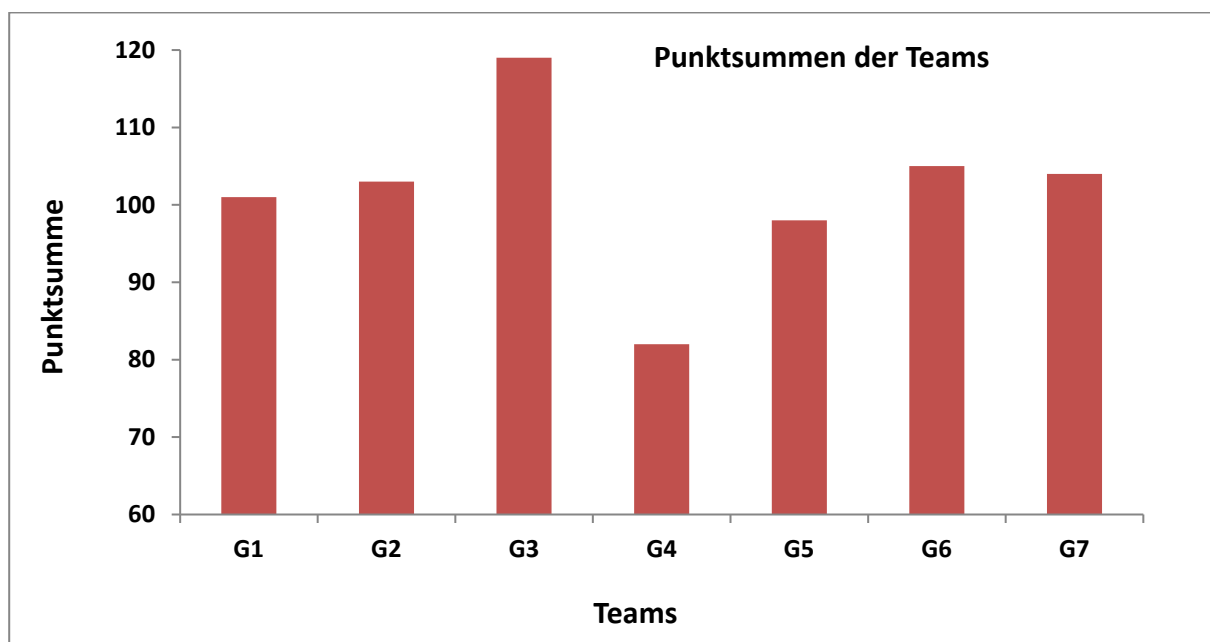


Abbildung 5: Punktschichte der Bewertungen für die einzelnen Teams für Förderrunde 1 (Online-Bewertung)

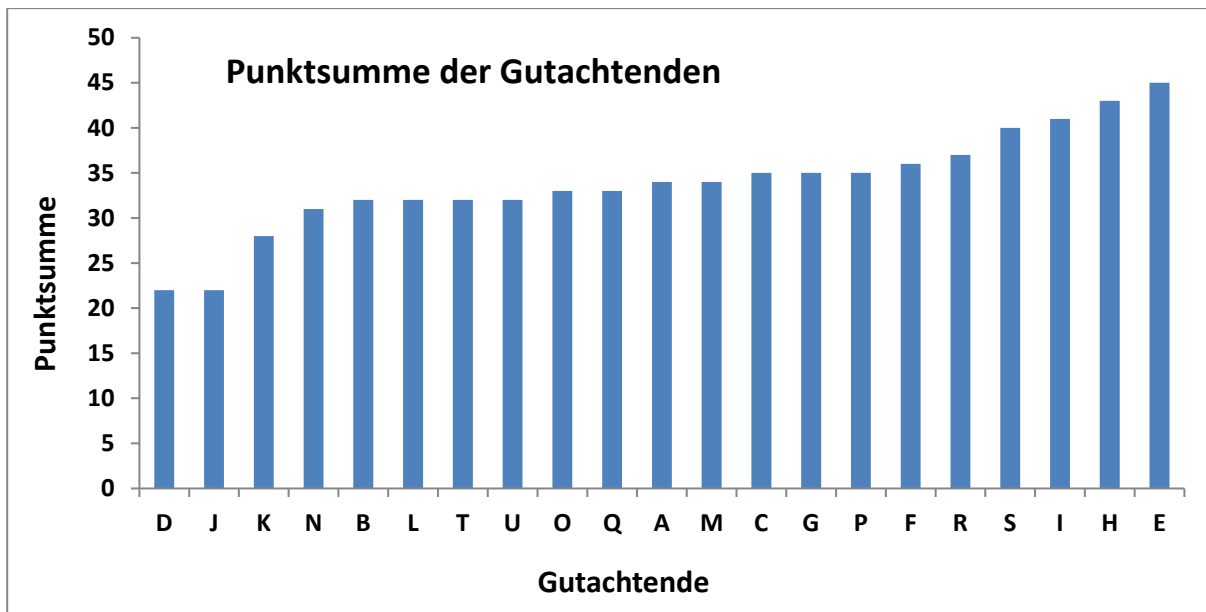


Abbildung 6: Punktsumme der Bewertungen für die einzelnen Gutachtenden für Förderrunde 1 (Online-Bewertung)

Die Berücksichtigung der Qualität der Anträge gelingt, wenn wir die Unterschiedlichkeit der Beurteilungen innerhalb der Gutachterteams betrachten. Veranschaulicht wird dies in Abbildung 7, in der die Punktsummen der individuellen Gutachtenden zusätzlich mit der Zuordnung zu den Teams zu sehen sind. Da die Gutachtenden innerhalb eines Teams mit identischen Förderanträgen konfrontiert waren, sind die Unterschiede in den Beurteilungen nur durch unterschiedliche Beurteilungskriterien der Gutachtenden (Gutachterbias) zu erklären. Besonders deutlich werden Unterschiede in den Teams G2 und G4, während die Teams G1 und G5 sehr ausgeglichen wirken.

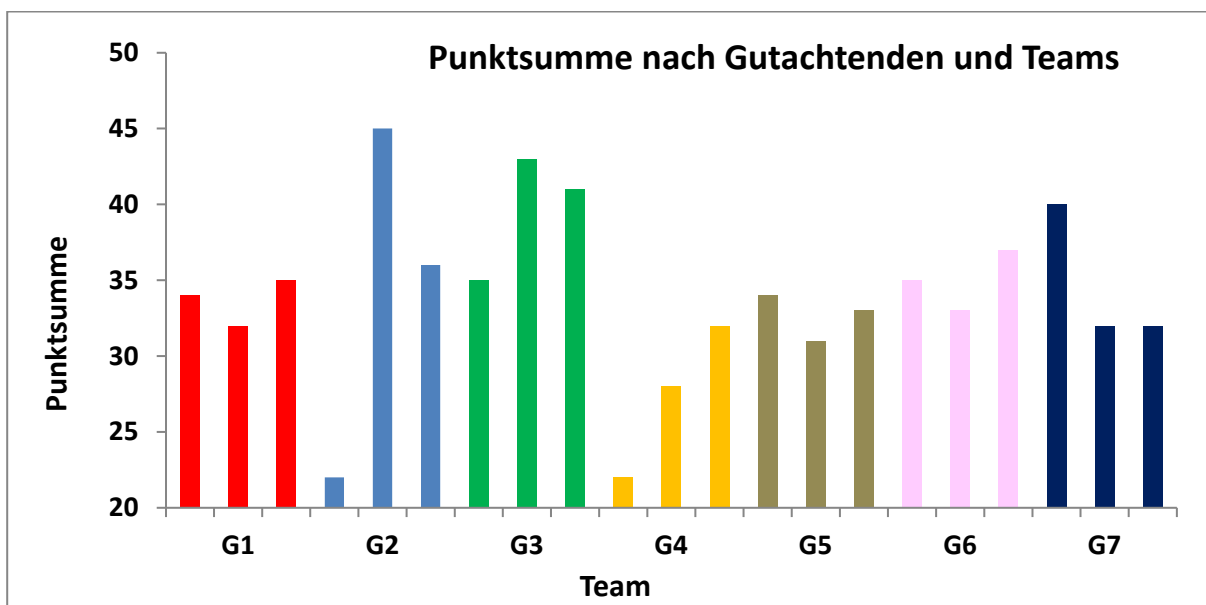


Abbildung 7: Punktsumme der Bewertungen für die einzelnen Gutachtenden mit Zuordnung zu den Teams für Förderrunde 1 (Online-Bewertung)

Betrachtet man die Beurteilungen zusätzlich in Abhängigkeit von der Gutachterperspektive (Abbildung 8), so ist zu erkennen, dass die Professorinnen und Professoren strenger urteilten als die Personen aus der Hochschuldidaktik und die Studierenden. Das könnte daran liegen, dass die professoralen



Gutachtenden mehr Erfahrung mit Begutachtungsprozessen haben und die anderen Gruppen aus einer größeren Unsicherheit heraus eher positivere Urteile vergeben.

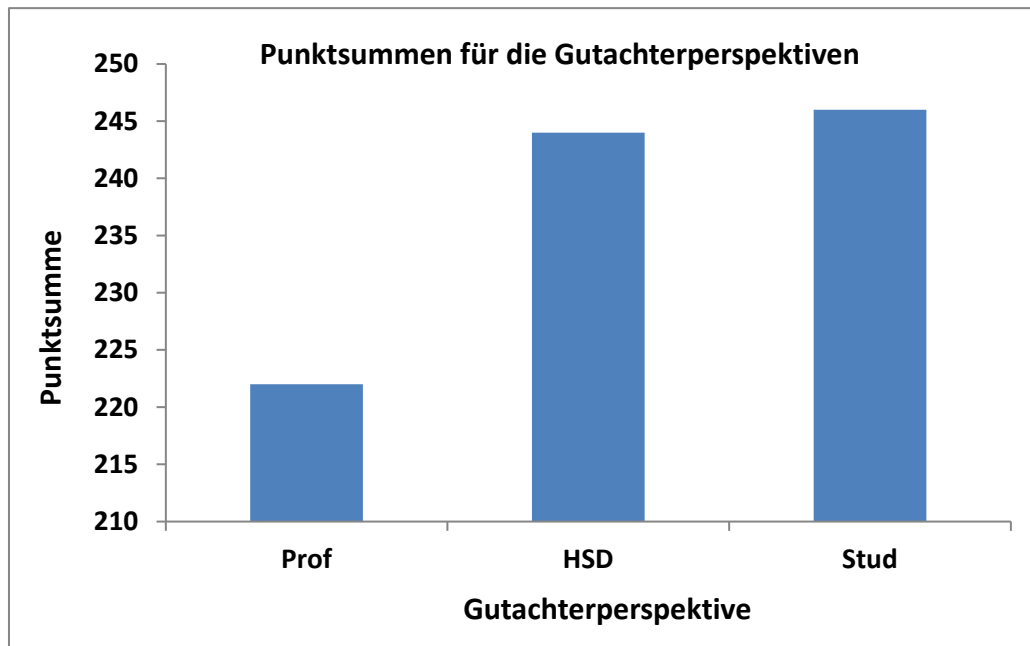


Abbildung 8: Punktsumme der Bewertungen für die unterschiedlichen Gutachterperspektiven für Förderrunde 1
Prof = professorale Gutachtende
HSD = Hochschuldidaktik
Stud = Studierende

6 Bewertungsanpassung durch die Gutachtersitzung

Nun zur zweiten Phase des Begutachtungsverfahrens: Basierend auf den Ergebnissen der Online-Bewertung trafen sich die Gutachtenden in der Großgruppe von 21 Personen und in den drei Personen starken Gutachterteams. Es wurde erwartet, dass sich diese Verfahrensroutine in der Qualität der Bewertungen (die hier nicht direkt messbar ist) niederschlägt, aber zumindest in einer Anpassung der Wertungen am Ende der Gutachtersitzung sichtbar wird. (Ansonsten würde ein rigides Beharren auf gutachterlichen Positionen vorliegen.)

Zunächst ein Blick auf die beurteilten Projektanträge: In Abbildung 9 ist zu sehen, inwieweit die Bewertungen korrigiert wurden. Bei 90 von insgesamt 154 Anträgen haben die Gutachtenden ihre Bewertungen nicht geändert, bei 64 Anträgen gab es Änderungen. Dabei konnte es zu einer Aufwertung oder einer Abwertung kommen. Bei den einzelnen Anträgen konnten sich ein, zwei oder alle drei Gutachtende eines Teams für eine Auf- oder Abwertung entscheiden. Die roten Balken in der Abbildung zeigen die Aufwertungen, die blauen die Abwertungen (mit ein, zwei oder drei Wertungsänderungen je Antrag).

In der Regel wurde die Bewertung dabei um jeweils einen Punkt angepasst. Eine Änderung um zwei oder drei Punkte kam nur in drei Fällen vor.

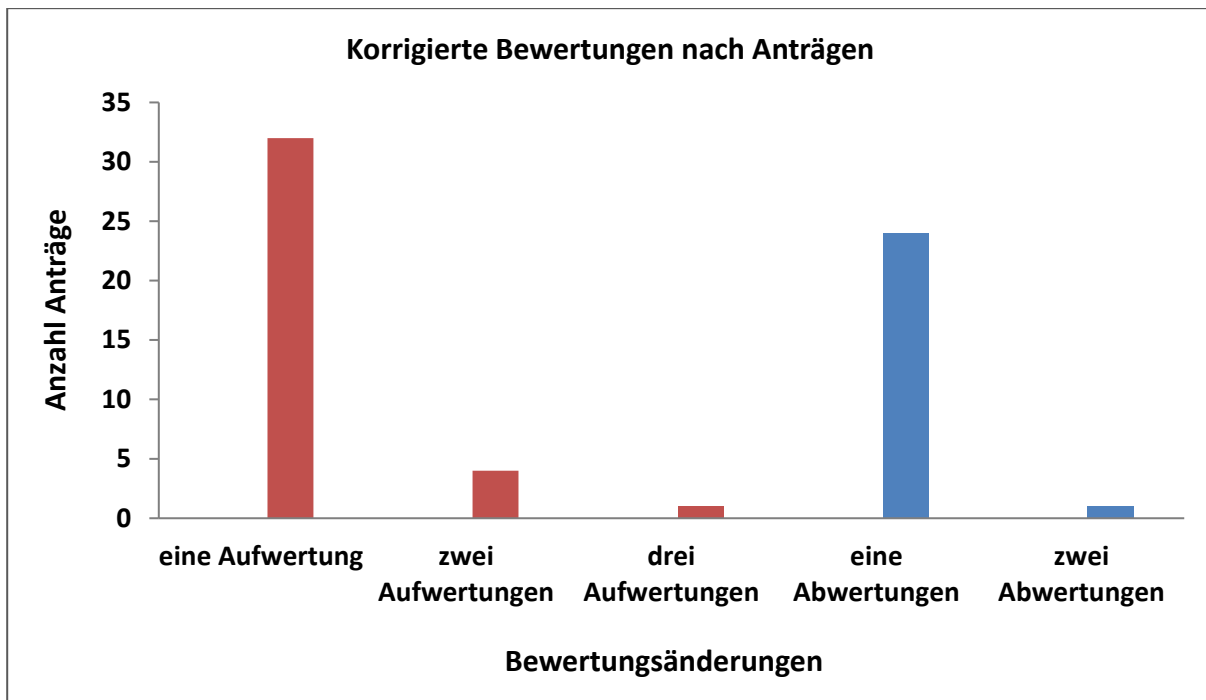


Abbildung 9: Änderung der Bewertung in der Gutachtersitzung im Vergleich zur Online-Bewertung. Aufwertungen sind als rote Balken dargestellt, Abwertungen als blaue. Eine, zwei oder drei Bewertungsänderungen bezieht sich auf die Anzahl der Personen im Team, die ihre Bewertung geändert haben.

Die nachfolgende Abbildung 10 zeigt die Bewertungsänderungen in Bezug auf die Gutachterteams. Die stark positiven Änderungen bei den Teams 2 und 6 stehen nicht im Zusammenhang mit einer besonders negativen Online-Bewertung bei diesen Teams. Auch die starke Abwertung in Team 7 lässt sich nicht durch einen großen Anpassungsbedarf erklären.

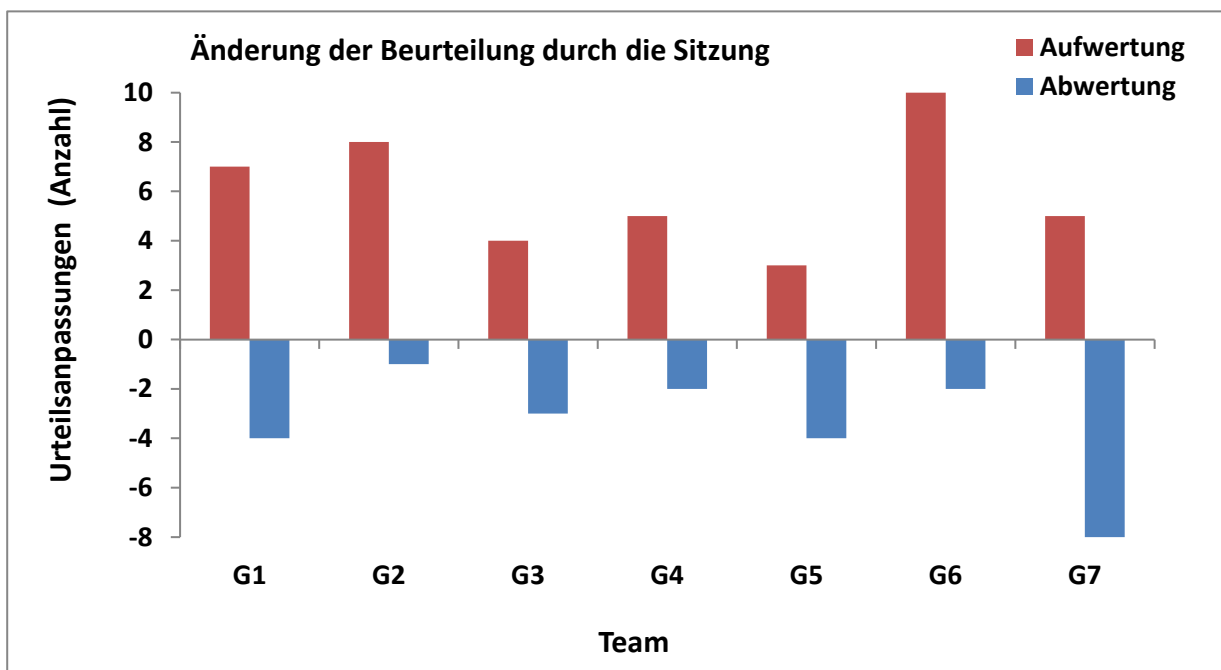


Abbildung 10: Änderung der Beurteilungen im Rahmen der Gutachtersitzung. Rot dargestellt sind die positiveren Bewertungen (Aufwertung), blau die negativeren (Abwertung).



Betrachtet man die Anpassungen der einzelnen Gutachtenden in den Teams (Abbildung 11), so zeigt sich ein uneinheitliches Bild: 14 Gutachtende passen ihre Urteile sowohl durch Auf- als auch durch Abwertungen an. Viermal finden nur Aufwertungen statt, zweimal nur Abwertungen. Nur eine Person (im Team G3) hat ihre Beurteilungen nicht geändert.

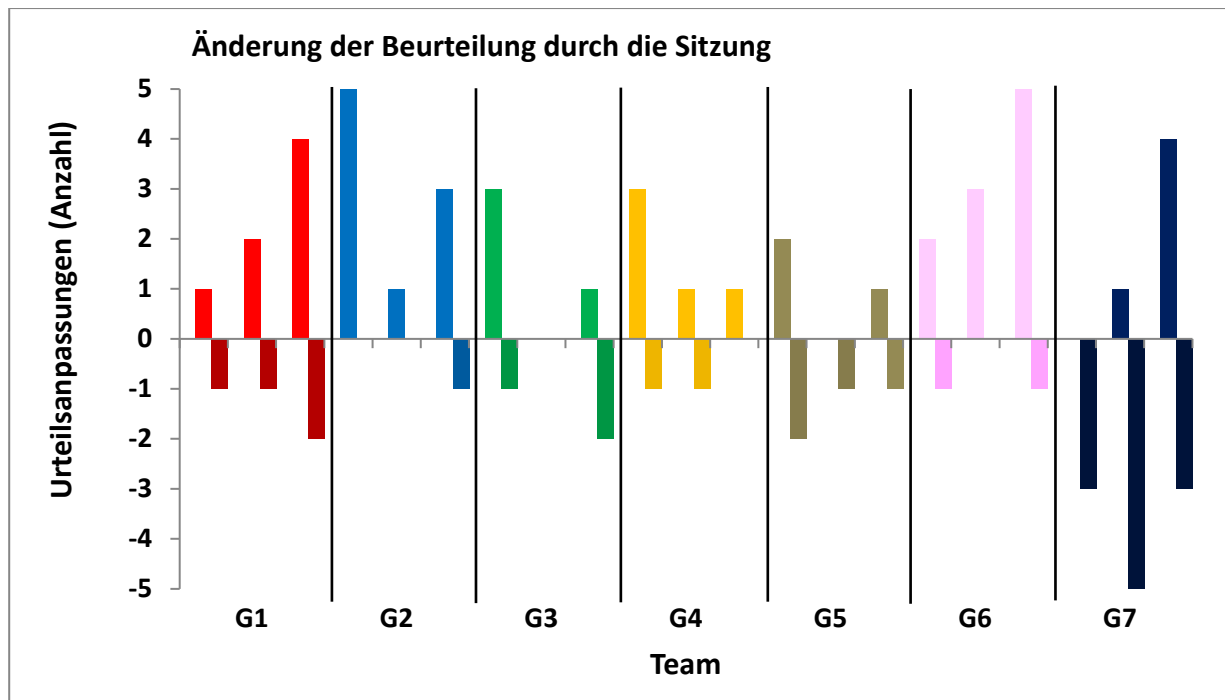


Abbildung 11: Änderung der Beurteilungen der Teams im Rahmen der Gutachtersitzung. Aufwertungen sind als positive Werte dargestellt, Abwertungen als negative. Zwei dicht nebeneinander liegende Säulen nach oben und unten gehören jeweils zu einer Person.

Inwieweit ist die Anpassung der Beurteilung abhängig von der Gutachterperspektive (professoral-hochschuldidaktisch-studentisch)? Wie in Abbildung 12 zu sehen ist, haben die Gutachtenden der Hochschuldidaktik ihre Bewertungen am wenigsten angepasst. Professorale und studentische Gutachtende haben umfangreichere Korrekturen vorgenommen, insbesondere bei der Aufwertung von Beurteilungen.

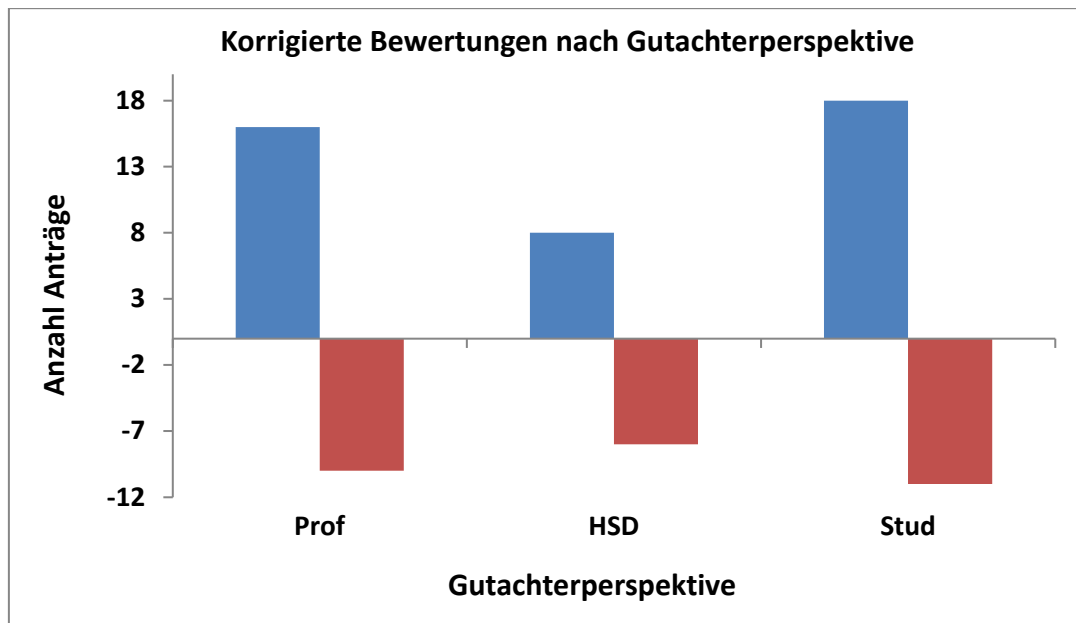


Abbildung 12: Änderung der Beurteilungen im Rahmen der Gutachtersitzung in Abhängigkeit von der Gutachterperspektive. Rot dargestellt sind die positiveren Bewertungen (Aufwertung), blau die negativeren (Abwertung).

Prof = professorale Gutachtende

HSD = Hochschuldidaktik

Stud = Studierende

7 Fazit

Mit dem zweistufigen Begutachtungsverfahren bei der Auswahl von Förderanträgen bot sich die Gelegenheit, einen kleinen **Einblick in die Entscheidungsprozesse** zu werfen, die in gutachtergestützten Verfahren im Hintergrund ablaufen. Eine erste Urteilsbildung erfolgt auf individueller Ebene, und es besteht die Möglichkeit der Anpassung dieses Urteils nach einem fachlichen Austausch innerhalb eines kleinen Teams von Spezialisten, die noch dazu ihre verschiedenen Perspektiven in die Beurteilung einbringen.

Dieses Vorgehen ist deutlich **aufwändiger** als das Einholen singulärer Urteile oder mehrerer individueller Urteile ohne weitere Abstimmung in einer Gruppe. Wie an den Daten zu sehen ist, gelingt es durch die Gruppenprozesse, Einschätzungen und Urteile zu verändern und damit den Auswahlprozess zielgerichtet voranzutreiben. Das „bessere“ Urteile sowohl für einzelne Anträge als auch für die gesamte Projektauswahl die Folge sind, kann nicht nachgewiesen werden, erscheint aber plausibel.

Lohnt sich der erhöhte Aufwand für das Verfahren? Sowohl der administrative Aufwand bei der Agentur als auch der zeitliche Aufwand bei den Gutachtenden steigt an. Außerdem ist der finanzielle Aufwand höher, insbesondere wenn sich die Großgruppe und die Teams an einem zentralen Ort treffen, verbunden mit einer Reisetätigkeit, Übernachtung und Verpflegung. Aber selbst in diesem Fall lagen die Kosten für das zweistufige Auswahlverfahren bei unter einem Prozent der gesamten Fördersumme (pro Förderrunde). Durch Videokonferenzen ließen sich die Kosten noch einmal deutlich reduzieren. In der letzten Förderrunde wurden Videokonferenzen explizit auf Wunsch der Gutachtergruppe eingesetzt, für die die Erfahrungen mit vorangegangenen videokonferenzbasierten Abstimmungen sehr positiv waren. Dabei ist zu erwähnen, dass sich die meisten Personen in der Gutachtergruppe bereits



durch vorangegangene Runden kannten und in ihrer Arbeit eingespielt waren. Für komplett neue Gruppenzusammensetzungen könnte ein „Live“-Treffen durchaus von Vorteil sein, da das Teambuilding über rein fachliche Gespräche hinausgeht.

Bei der **Auswahl von Projektanträgen** muss im Einzelfall entschieden werden, ob sich der Aufwand lohnt. Das ist abhängig von der Bedeutsamkeit der Auswahlentscheidung, den Kosten, die durch Fehlentscheidungen entstehen, aber auch von der Fördersumme und Laufzeit der Förderung. Für die **Gutachtenden** scheint sich der Aufwand überwiegend gelohnt zu haben, da sich viele entschieden haben, in mehreren aufeinanderfolgenden Runden mitzuwirken. Von den Gutachtenden wurde in Feedbackgesprächen auch hervorgehoben, dass der Austausch in den Gruppen sehr wertvoll sei und ein guter Einblick in die Innovationsprojekte der Hochschulen gegeben werde. Damit verbunden waren auch Anregungen für die Entwicklung der eigenen Lehre. (Systematisch erhoben wurde die Motivation zur Teilnahme bei den Gutachtenden jedoch nicht.)

Es ist zu erwarten, dass auch die **Antragsteller** der Projekte von dem Vorgehen profitieren, da sich ihre Investition in den Förderantrag eher lohnt, als wenn eine reine Zufallsauswahl von formal geeigneten Projekten erfolgt oder nur eine oberflächliche Beurteilung anhand formaler Kriterien vorgenommen wird. Dies wäre vergleichbar mit der ersten Stufe der Auswahl, der Online-Bewertung. Ohne die zweite Stufe der Gutachtersitzung wäre es zudem nicht möglich gewesen, den Antragstellern (insbesondere den abgelehnten) eine abgestimmte inhaltliche Stellungnahme zu ihrem Projekt und ihrem Antrag zu geben und ihnen damit die Chance zu eröffnen, bei der nachfolgenden Ausschreibung erfolgreich zu sein. Die Begutachtung ist somit Teil eines lernenden Systems.

Sehr gut bewährt hat sich im Verfahren die **Diversität** in der Zusammensetzung der Gutachterteams: Der Einsatz der unterschiedlichen Perspektiven *professorale Lehre – Hochschuldidaktische Expertise – Abnehmerseite Studierende* wurde von der Gutachtergruppe durchweg positiv beurteilt. Auffällig war, dass im Begutachtungsprozess, insbesondere in der Abstimmung im Team, keine Statusunterschiede zwischen den Personen aus den verschiedenen Perspektiven zu erkennen waren und auch nicht aus den Gruppen berichtet wurden. Die Personen haben sich gegenseitig als Expertinnen und Experten für Studium und Lehre wahrgenommen und respektiert.

Was fehlt? Es handelt sich im vorliegenden Fall um eine reine Post-Hoc-Analyse des Begutachtungsverfahrens und nicht um eine geplante Studie. Sollte sich erneut die Gelegenheit ergeben, ein ähnlich aufwändiges Begutachtungsverfahren durchzuführen, so sollte die Verfahrensevaluation bereits in der Planungsphase berücksichtigt werden. Es wäre interessant zu erfahren, wie die Urteilsbildung aus Sicht der Gutachtenden erfolgt und wie insbesondere die Anpassung des eigenen Urteils im Gruppenprozess bewertet wird. Systematische Befragungen der Gutachtenden und der beurteilten Antragsteller zur Zufriedenheit mit dem Verfahren sollten zudem vorgesehen werden, um die Akzeptanz des Verfahrens zusätzlich abzusichern und das Verfahren ggf. weiterentwickeln zu können.

Lassen sich die Erkenntnisse übertragen? Es gibt im Hochschulsektor eine Vielzahl von Begutachtungsverfahren mit Beteiligung einzelner gutachtender Personen und die Einbettung von Individuen in Gutachtergruppen. Für bedeutsame und kostenintensive Fördermaßnahmen, bei denen man Fehlerurteile möglichst klein halten möchte, lohnt es sich, das aufwändige zweistufige Verfahren in Erwägung zu ziehen. Aber auch bei anderen Qualitätsbewertungen, z. B. dem alltäglichen Geschäft der Begutachtungen in Akkreditierungsverfahren (Programmakkreditierungen und Systemakkreditierungen) können die Erkenntnisse aus diesem Verfahren herangezogen werden. In der Regel liegt dort der Fokus



auf der Gruppenbewertung bei Vor-Ort-Gesprächen in den Hochschulen. Aber vielleicht lohnt es sich, individuelle Vorabbewertungen auf Basis eines Selbstberichts der Hochschule stärker strukturell im Verfahren zu verankern. Agenturen verwenden dieses Vorabstellungsverfahren vor den Vor-Ort-Gesprächen bereits in einigen Fällen, jedoch nicht immer als verbindlichen und standardisierten Prozess. Rechtsverordnungen zur Akkreditierung machen dazu keine Vorgaben. Durch die Einbindung dieser Phase ließen sich jedoch systematisch mehr Informationen zum Beurteilungsgegenstand im Vorfeld der Vor-Ort-Gespräche sammeln und den Gruppenprozess der Begutachtung vorbereiten. Die Entwicklung vom Einzel- zum Gruppenurteil würden den Gutachtenden zumindest explizit verdeutlicht werden. Anders als bei der Auswahl von Förderanträgen werden in Akkreditierungsbegutachtungen in den Vor-Ort-Gesprächen noch neue Informationen hinzugewonnen; die Förderanträge bieten jedoch ein stabile Informationslage. Der Mehrwert einer expliziten oder strukturierten individuellen Erstbewertung scheint aber dennoch gegeben.



Anhang: Online-Beurteilungsbogen

Beurteilung von Projektanträgen INNOVATION PLUS 2023/24 (Runde 5)

Förderprogramm für niedersächsische Hochschulen

Liebe Gutachterin, lieber Gutachter,

der Fragebogen dient der Beurteilung von Projektanträgen im Rahmen des Förderprogramms „Innovation plus“ des Landes Niedersachsen (Runde 5). Ziel der Förderung ist es, Freiräume für Lehrende zur Entwicklung von innovativen Lehr- und Lernkonzepten zu schaffen. Dabei wird insbesondere die Weiterentwicklung von Elementen von Modulen oder ganzen Modulen in den Fokus genommen. Das Programm versteht sich als Impulsgeber und Anschubfinanzierung für Lehrkonzepte, die bei Erfolg längerfristig Teil des Lehrangebots sein werden. Anträge können von allen hauptberuflich Lehrenden sowie von Gruppe von Lehrenden gestellt werden. *Die aktuelle Förderrunde soll sich in besonderer Weise der Studieneingangsphase widmen und hierbei einen oder mehrere der folgenden Aspekte in den Blick nehmen:*

- *Entwicklung von Modulen im ersten Semester zur Verbesserung der Studierfähigkeit*
- *Entwicklung oder Implementierung neuer Lehrformen für »Massenveranstaltungen« « in den ersten Semestern von grundständigen Studiengängen*
- *Einführung neuer hochschuldidaktischer Konzepte und Maßnahmen*
- *Stärkung des Praxisbezugs in grundständigen Studiengängen*
- *Verbesserung der Lehre durch Digitalisierung*
- *Integration von Forschung in die Lehre*
- *Entwicklung von Konzepten zum Umgang mit zunehmender Diversität der Studierendekohorten*
- *Förderung der Kommunikation und Zusammenarbeit von Studierenden (Kollaboratives Lernen)*

Darüber hinaus sollen auch die Prüfungsformen beachtet werden. Die Ausschreibung besagt: "Bei den Projektanträgen sind insbesondere die Entwicklung und Evaluation von Methoden zur Überprüfung des Kompetenzzuwachses der Studierenden zu berücksichtigen. Hierfür sind adäquate kompetenz- und lernzielorientierte Prüfungsformate einzusetzen. Es sollte beschrieben werden, welche Möglichkeiten der Prüfungsgestaltung hierfür den Projektverantwortlichen zur Verfügung stehen."

Außerdem soll im Antrag ein Konzept skizziert werden, wie die Pilotmodule bei positiver Evaluation nach Abschluss der Förderung in die Curricula der Studiengänge integriert werden können.

Mit dem vorliegenden Beurteilungsbogen sollen die Gutachterinnen und Gutachter die Qualität der Projektanträge beurteilen und abschließend eine Empfehlung zur Förderung des Projekts abgeben. Jedes Mitglied der Gutachtergruppe soll jeweils einen Beurteilungsbogen für die ihr/ihm zugewiesenen Projekte ausfüllen.

Hinweise zum Ausfüllen

Das überwiegende Antwortformat ist eine 4-stufige Skala, mit der ein Ausprägungsmerkmal oder eine Güte erfasst wird. Hinter jedem Item-Block (Fragen oder Aussagen) findet sich zudem auch ein Freitextfeld für Kommentare, über das die Beurteilungen bei Bedarf erläutert werden können.

Die dreistellige Antragsnummer (Frage 2) benötigen wir, um die Beurteilungen später filtern zu können. Sie befindet sich in roter Schrift auf dem ersten Blatt des Antrags und im Namen der PDF-Datei. (Es sollten immer genau drei Ziffern eingegeben werden, sonst funktioniert die Filterung später nicht.)

Hinweise zur Beurteilung

Zur Beurteilung soll primär das Item 7 (Gesamtbeurteilung) herangezogen werden.

Die Items 5 und 6 wurden direkt aus der Ausschreibung abgeleitet und sollen Sie bei Ihrer Bewertung unterstützen. Außerdem sollen diese Angaben herangezogen werden, wenn sich stark unterschiedliche Einschätzungen bei den Gutachterinnen un

ALLGEMEINE ANGABEN

* 1. Name d. Gutachtenden (bitte aus der Liste auswählen)

* 2. Dreistellige Antragsnummer (z.B.: 001 - bitte aus dem Antrag übernehmen)

* 3. Name der antragstellenden Hochschule (bitte aus der Liste auswählen)

* 4. Name des Antragstellers (bitte aus dem Namen der PDF-Datei übernehmen)

BEURTEILUNG

* 5. Themenfelder

(Beurteilung, inwieweit die Themenfelder berücksichtigt werden)

	sehr stark	eher stark	eher schwach	sehr schwach / gar nicht	nicht beurteilbar, zu wenige Informationen
Das Projekt widmet sich der Studieneingangsphase	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Module des 1. Semesters werden genutzt, um die Studierfähigkeit zu verbessern	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Neue Lehr- und Lernformen für Massenveranstaltungen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Stärkung des Praxisbezugs in grundständigen Studiengängen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbesserung der Lehre durch Digitalisierung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Integration von Forschung in die Lehre	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Umgang mit zunehmender Diversität	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Förderung der Kommunikation und Zusammenarbeit von Studierenden (kollaboratives Lernen)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Erläuterungen (bei Bedarf)

*** 6. Auswahlkriterien**

(Beurteilung in Bezug auf die jeweilige Fachkultur)

	sehr hoch	eher hoch	eher gering	sehr gering / gar nicht	nicht beurteilbar, zu wenige Informationen
Beitrag zur Unterstützung der Studierenden beim Studienbeginn	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beitrag des Projekts zur Verbesserung der Lehr- und Prüfungsqualität	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Innovationscharakter der Projektidee (bezogen auf die Fachkultur)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Entwicklungspotenzial der Projektidee (bezogen auf die Fachkultur)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nachhaltigkeit der Projektidee über den Förderzeitraum hinaus	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Berücksichtigung kompetenz- und lernzielorientierter Prüfungsformate	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Realisierbarkeit unter Berücksichtigung der verfügbaren Ressourcen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Erfolgskontrolle durch geeignete Qualitätssicherungsmaßnahmen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beteiligung hochschulinterner Expertise für die Sicherung und Entwicklung der Qualität in Studium und Lehre	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Erläuterungen (bei Bedarf)

*** 7. Wurde ein Konzept skizziert, wie die Pilotmodule bei positiver Evaluation nach Abschluss der Förderung in die Curricula der Studiengänge integriert werden können?**

- ja, es wurde ein Konzept dargestellt
- nein, es wurde kein Konzept dargestellt

Erläuterungen (bei Bedarf):

8. Falls ja, erscheint das dargestellte Konzept erfolgversprechend und tragfähig?

- ja, das Konzept erscheint erfolgversprechend und tragfähig
- nein, das Konzept erscheint nicht erfolgversprechend und tragfähig

Erläuterungen (bei Bedarf):

FAZIT

* 9. Gesamtbeurteilung (Förderempfehlung)

Auf Basis der oben beurteilten Kriterien soll eine Gesamtbeurteilung für das Projekt abgeleitet werden. Für die Gesamtbeurteilung des Projektantrags sind vier gestufte Kategorien vorgesehen (s.u.). Jeder Kategorie wird ein Punktwert zugeordnet, so dass wir für jeden Antrag (der jeweils von drei Personen begutachtet wird) eine Punktschätzung ermitteln können. Dieser Score hat sich bei den Beurteilungen der letzten Förderrunde als sehr hilfreich erwiesen.

Anmerkung: Die Kategorie „ist nur eingeschränkt zur Förderung geeignet“ bedeutet, dass die Qualität des Projekts/Antrags zwischen den Beurteilungen „keine Förderung“ und „Förderung“ einzuschätzen ist. Eine „vorrangige Förderung“ sollte nur für besonders positive Projekte/Anträge gewählt werden.

Das Projekt ...

- ... sollte **nicht** gefördert werden (0 Punkte)
- ... das Projekt ist nur **eingeschränkt** zur Förderung geeignet (1 Punkt)
- ... sollte **gefördert** werden (2 Punkte)
- ... sollte **vorrangig** gefördert werden (3 Punkte)

Erläuterung: Die Beurteilung sollte kurz begründet werden, so dass die verschiedenen Positionen bei abweichenden Meinungen in der Gutachtergruppe transparent werden.

Zudem kann der Text später für ein Feedback an die Antragsteller herangezogen werden.



ZEvA

Die ZEvA führte in den Jahren 2018 bis 2022 im Auftrag des Niedersächsischen Ministerium für Wissenschaft und Kultur ein gutachterbasiertes Auswahlverfahren von Förderanträgen niedersächsischer Hochschulen durch. Dafür wurde ein zweistufiges Verfahren gewählt: Aufbauend auf individuelle Gutachterbewertung über einen Online-Fragebogen wurde eine finale Beurteilung durch eine große Gutachtergruppe und mehrere kleine Gutachterteams vorgenommen. Durch dieses gestufte Verfahren, das mit intensiven Abstimmungsprozessen in der Großgruppe und in den Teams verbunden war, konnte ein qualitativ hochwertiger, aber auch aufwändiger Auswahlprozess etabliert werden. Die Begutachtungen waren verbunden mit iterativen Prozessen von individuellen Urteilen zu Gruppenurteilen und führten insgesamt zu einer hohen Zufriedenheit bei den Gutachterinnen und Gutachtern.