Deutscher Juristinnenbund

Policy Paper zur geschlechtergerechten Bestenauslese im öffentlichen Dienst

05.09.2025

Inhalt

A. Einleitung 1
B. Herausforderungen 2
I. Mängel bei der Gleichbehandlung von Frauen und Männern 2
II. Fehlen einer sachlichen Grundlage für Gendergerechtigkeit 4
1. Mängel im Beurteilungsverfahren 4
2. Mängel bei den Beurteilungskriterien 7
C. Lösungen 10
I. Vereinheitlichung der Beurteilungsregelungen 10
II. Sachliche Beurteilungsgrundlage zur Schaffung von Gendergerechtigkeit 11
1. Sachlichkeit im Beurteilungsverfahren 11
2. Sachlichkeit bei den Beurteilungskriterien 13
D. Zusammenfassung 14

A. Einleitung

Die neue Bundesregierung hat sich in ihrem Koalitionsvertrag „Verantwortung für Deutschland“ zum Ziel gesetzt, mehr Frauen in Führungspositionen zu bringen (Zeile 1843).[1] Bei allen Fortschritten sprechen die Zahlen deutlich gegen eine tatsächliche Gleichstellung von Frauen und Männern im öffentlichen Dienst.

Das Statistische Bundesamt erstellt jährlich auf der Grundlage der zuletzt durch das Zweite Führungspositionengesetz (FüPoG II) geänderten Gleichstellungsstatistikverordnung[2] den Gleichstellungsindex zur Gleichstellung von Frauen und Männern in den obersten Bundesbehörden. Diesem Gleichstellungsindex ist zu entnehmen, dass der Anteil von Frauen in Führungspositionen zum Stichtag 30. Juni 2024 nur bei 44,3 % lag. Er ist damit noch unrealistisch weit von dem in § 1 Abs. 2 S. 3 Bundesgleichstellungsgesetz (BGleiG) gesetzten Ziel entfernt, die „gleichberechtigte Teilhabe“ von Frauen und Männern bis Ende 2025 zu erreichen. Diese „gleichberechtigte Teilhabe“ kann nach § 13 Abs. 2 S. 1 Nr. 1 BGleiG angenommen werden, sobald „Führungspositionen, in denen Frauen bisher unterrepräsentiert waren, mit annähernd numerischer Gleichheit mit Frauen und Männern besetzt werden“. Hierfür stehen die Chancen auch deshalb gering, weil das Statistische Bundesamt in seinem im Jahr 2025 veröffentlichten Gleichstellungsindex „nur wenig Fortschritt“ im Vergleich zum Vorjahr feststellt.

Die festzustellende Unterrepräsentanz von Frauen in Führungspositionen lässt sich auch nicht mit einem geringeren Anteil von Frauen im höheren Dienst allgemein erklären, denn nur beim Bundesministerium für Gesundheit (BMG) liegt der Anteil von Frauen und Männern im höheren Dienst mit dem der beiden Geschlechter in Führungspositionen gleichauf und nur bei den relativ kleinen Behörden der Bundesbeauftragten für den Datenschutz und die Informationsfreiheit (BfDI) und dem Bundesrat liegt der Anteil der Frauen in Führungspositionen sogar über dem der Frauen im höheren Dienst insgesamt. Ansonsten überwiegt der Anteil der Männer in Führungspositionen im Verhältnis zu ihrem Anteil im höheren Dienst und liegt oftmals sogar eklatant über dem entsprechenden Verhältnis für Frauen.

Im direkten Zusammenhang mit dem Ziel, mehr Frauen in Führungspositionen zu bringen, beschränkt sich der Koalitionsvertrag auf die Nennung von Aspekten der Vereinbarkeit von Beruf und Care-Verpflichtungen, „flexiblere Arbeitszeitmodelle, bessere Möglichkeiten für Führen in Teilzeit“. Vielleicht können sich Frauen im öffentlichen Dienst auch mitgemeint fühlen, wenn sich die Bundesregierung noch etwas allgemein „eine bessere Abbildung der Vielfalt unserer Gesellschaft in der öffentlichen Verwaltung“ als Ziel gesetzt hat.[3] Damit Frauen in Führungspositionen gelangen, werden derartige Maßnahmen allerdings nicht ausreichen. Das gilt auch für die Förderung von Führen in Teilzeit, zu dem das BMFSFJ bereits in der vergangenen Legislaturperiode einen vielbeachteten Handlungsleitfaden[4] herausgebracht hat. Denn Geschlechtergerechtigkeit muss ab Beginn der Laufbahn bestehen und darf nicht erst beim Sprung in die Führungsposition einsetzen.

Ausgehend von Art. 33 Abs. 2 Grundgesetz (GG) bleibt es bei allen berechtigten Begleitmaßnahmen dabei, dass entscheidend für das berufliche Fortkommen von Beamt*innen die dienstliche Beurteilung

---

[1] Koalitionsvertrag „Verantwortung für Deutschland“ zwischen CDU, CSU und SPD, 21. Legislaturperiode (https://www.koalitionsvertrag2025.de/sites/www.koalitionsvertrag2025.de/files/koav_2025.pdf).

[2] GleiStatV.pdf. Eine weitere Rechtsgrundlage für eine Erfassung des „geschlechtsspezifischen Entgeltgefälles“ wird der nach Art. 34 Abs. 1 S. 1 bis zum 7. Juni 2026 in nationales Recht umzusetzende Art. 9 der EU-Entgelttransparenzrichtlinie (Richtlinie - 2023/970 - EN - EUR-Lex) sein.

[3] Koalitionsvertrag „Verantwortung für Deutschland“ zwischen CDU, CSU und SPD, 21. Legislaturperiode (https://www.koalitionsvertrag2025.de/sites/www.koalitionsvertrag2025.de/files/koav_2025.pdf).

[4] Abrufbar unter https://www.bmbfsfj.bund.de/resource/blob/241356/9a2a5f5a9bed25e7f9d5f6bf07895ace/handlungsleitfaden-data.pdf.

ist.[5] So wichtig die dienstliche Beurteilung ist, so offen ist sie für subjektive Bewertungen bis hin zur Manipulation. Während sich schriftliche Prüfungen anonymisiert durchführen lassen, geht es bei der Beurteilung von Leistung und Befähigung nach Art. 33 Abs. 2 GG gerade darum, dass eine beurteilende Person sich zu einer ihr persönlich bekannten Person äußert. Entsprechend können Vorurteile, persönliche Motivationen und andere sachfremde Faktoren die dienstliche Beurteilung beeinflussen. Kernanliegen für Chancengleichheit von Frauen und Männern müssen im Beurteilungswesen also Gleichbehandlung und Transparenz sowie sachgerechte Beurteilungsverfahren und -kriterien sein.

Dieses Policy Paper untersucht zunächst (unter B), an welchen Voraussetzungen für eine geschlechtergerechte Beurteilungspraxis es fehlt. In einem weiteren Teil (unter C) werden dann Lösungsansätze zur Beseitigung der beschriebenen Defizite unter Berücksichtigung der Zielsetzungen des Koalitionsvertrags ausgelotet, bevor das Ergebnis (unter D) zusammengefasst wird. Damit will der djb der Bundesregierung praktische Vorschläge unterbreiten und einen Beitrag zum Diskurs im Parlament und im außerparlamentarischen Raum zur Umsetzung eines geschlechtergerechten Beurteilungswesens leisten.

B. Herausforderungen

I. Mängel bei der Gleichbehandlung von Frauen und Männern

Voraussetzung für jede Gleichbehandlung nach Art. 3 GG, und damit auch für die von Frauen und Männern, ist, dass auf vergleichbare Fälle die gleichen Standards angewandt werden. Was wie eine juristische Binsenweisheit klingt, berührt ein grundsätzliches Problem des Beurteilungswesens. Es gelten nämlich schon auf Bundesebene eine Vielzahl unterschiedlicher Regelungen, von Ländern und Kommunen ganz zu schweigen. Auf der Grundlage des Ressortprinzips nach Art. 65 S. 2 GG haben alle Ministerien ihre eigenen Beurteilungsrichtlinien gegeben, denn die Bundesressorts sind interessiert, „sich Gestaltungsspielräume auf Ressortebene zu erhalten“.[6] Innerhalb der Geschäftsbereiche gibt es darüber hinaus noch behördenspezifische Beurteilungsrichtlinien. Da sämtliche dieser Richtlinien auf der Grundlage der §§ 48 bis 50 der Bundeslaufbahnverordnung (BLV)[7] als behördeninternes Recht erlassen werden,[8] müssen sie – anders als Rechtsnormen – nicht öffentlich zugänglich gemacht werden.

Diese mangelnde Transparenz der Beurteilungssysteme steht einer Vergleichbarkeit zwischen den Anwendungsbereichen der Beurteilungsrichtlinien und damit einer Gleichbehandlung schon ganz grundsätzlich entgegen, da schlicht die nötigen Informationen über die Maßstäbe für einen Vergleich fehlen. So stand zum Beispiel für dieses Policy Paper nur eine veröffentlichte Beurteilungsrichtlinie zur Verfügung, die auf der Webseite von „FragDenStaat“ zu finden war.[9] Ansonsten haben sich die Autorinnen auf eine Reihe von Richtlinien und ihre Anlagen gestützt, die dem djb aus anderer Befassung bekannt, aber nicht öffentlich zugänglich sind.

Wo konkrete dienstliche Beurteilungen zum Vergleich vorliegen, scheitert eine Vergleichbarkeit meist schon daran, dass die verschiedenen Beurteilungsrichtlinien nutzen nämlich unterschiedliche Beurteilungsskalen. Zu den komplexeren Benotungsskalen dürfte folgende Regelung gehören: „Für die Bewertung stehen die Notenstufen A1 bis A4, X1 und X2 sowie B und C zur Verfügung.“ Alternative

[5] Z. B. BVerfG, 2 BvR 1558/16, Rn. 8 m.w.N.
[6] *Jürgen Lorse*, Die dienstliche Beurteilung von Frauen. Strukturelle Diskriminierung beseitigen, Bewusstsein verändern, Die öffentliche Verwaltung [DÖV] 2017, S. 455-464 (456).
[7] Die Bundeslaufbahnverordnung ist wiederum auf Grund mehrerer Vorschriften des Bundesbeamtengesetzes ergangen (s. Eingangsformel: https://www.gesetze-im-internet.de/blv_2009/eingangsformel.html).
[8] BVerwG 2 B 104.11, Rn. 5 stellt klar: „Beurteilungsrichtlinien sind keine Rechtsnormen“.
[9] https://fragdenstaat.de/anfrage/beurteilungsrichtlinien-der-beamtinnen-und-beamten-in-ihrem-zustaendigkeitsbereich/.

Lösungen mit Buchstaben von A bis F (mit A als Bestnote) oder Nummern von 1 bis 6 (mit 6 als Bestnote) für die Leistungsbewertung kommen sogar innerhalb eines Geschäftsbereichs vor.[10] Dem Einfallsreichtum an Buchstaben- und Nummernsystemen sind keine Grenzen gesetzt.

Außerdem ist die Vergleichbarkeit der dienstlichen Beurteilungen dadurch erschwert, dass es unterschiedliche Beurteilungszeitpunkte gibt. Auf Länderebene wird sogar innerhalb eines Beurteilungssystems auf Regelbeurteilungen völlig verzichtet,[11] oder es werden dem Worte nach Regelbeurteilungen vorgesehen, deren Beurteilungsrhythmus sich allerdings individuell (etwa nach Einstellungsdatum) bestimmt, so dass Beamt*innen nach einem persönlichen Regelrhythmus beurteilt werden, der mit dem ihrer Kolleg*innen nicht übereinstimmt.[12]

Zusätzlich zur unübersichtlichen Vielzahl von Beurteilungssystemen steht auch eine entsprechende Variantenbreite an Beurteilungskriterien einem transparenten, geschlechtergerechten Beurteilungswesen entgegen. Oftmals ist unklar, ob abweichende Begriffe doch vielleicht dasselbe bedeuten sollen, etwa bei den Kriterien Arbeitsverhalten und Leistungsverhalten, Arbeitsqualität und Arbeitsweise, soziales Verhalten und soziale Kompetenzen. Jedenfalls liegt es nahe, dass unterschiedliche Begriffe auch unterschiedlich ausgelegt werden und damit einer Vergleichbarkeit der auf diesen Beurteilungsregelungen erstellten dienstlichen Beurteilungen im Wege stehen.

Gleichbehandlung von Frauen und Männern im öffentlichen Dienst oder auch nur in der Bundesverwaltung lässt sich auch dann nicht ohne weiteres herstellen, wenn man die Möglichkeit von Ressortwechseln außer Acht lässt, obwohl die neue Bundesregierung solche Flexibilität eigentlich stärker nutzen will[13] und obwohl Ressortwechsel tatsächlich keine Seltenheit sind. Beschränkt man sich in der Betrachtung also auf den Anwendungsbereich einer einzelnen Beurteilungsrichtlinie, so ist eine Gleichbehandlung bereits förmlich in Frage gestellt, wenn zur Einhaltung von Richtwerten keine angemessenen Vergleichsgruppen gebildet werden können. So können zum Beispiel kleine Behörden mit nur wenigen Beamten keine ausreichend großen Vergleichsgruppen für die einzelnen Statusämter herstellen. Wann eine solche Gefahr der mangelnden Größe einer Vergleichsgruppe besteht, lässt sich nicht ohne weiteres feststellen. Das Bundesverwaltungsgericht hat zwar im Jahr 2005 schon einmal eine Vergleichsgruppenbildung für nicht groß genug befunden.[14] Im Jahr 2017 hat es dann aber betont, die Mindestgröße einer Vergleichsgruppe lasse sich „nicht allgemein, sondern nur für den konkreten Fall beantworten“.[15] Das Problem der Vergleichsgruppenbildung hat das Bundesverwaltungsgericht in derselben Entscheidung aus dem Jahr 2017 weiter verkompliziert, indem es betonte, die Vergleichsgruppe müsse auf eine für den „einzelnen Beurteiler [...] überschaubare Gruppe bezogen“ sein.[16] Selbst wenn man wechselnde Personalausstattungen unberücksichtigt lässt, ist es also schwierig zu gewährleisten, dass die Vergleichsgruppe in einem Statusamt zugleich groß genug ist, um einen Vergleich zu ermöglichen, und nicht zu groß ist, um für die beurteilende Person unübersichtlich zu sein.

Es ist daher oftmals – etwa wegen der Unterschiede zwischen Beurteilungssystemen, wegen zu kleiner oder für einen einzelnen Beurteilenden zu unübersichtlich großer Vergleichsgruppen – auch bei bestem Willen der Beurteilenden fragwürdig, ob Gleichbehandlung bei dienstlichen Beurteilungen überhaupt

---

[10] Auf Quellenangaben wird hier verzichtet, da diese Richtlinien, die dem djb aus verschiedenen Zufällen bekannt sind, ohnehin nicht öffentlich zugänglich sind.

[11] In Brandenburg gibt es keine Regelbeurteilungssysteme (Schnellenbach/Bodanowitz BeamtenR/Bodanowitz, 11. Aufl. 2024, § 11 Rn. 15). In Rheinland-Pfalz ist erst mit Wirkung zum 01.07.2025 ein Regelbeurteilungssystem (wieder) eingeführt worden, vgl. § 15 LaufbahnVO RP, geändert durch Verordnung vom 04.12.2024 (GVBl. S. 401).

[12] in Hamburg (Nr. 4.2 der Beurteilungsrichtlinie, auf S. 65 in der Broschüre zum Beurteilungswesen: https://www.hamburg.de/resource/blob/216634/d2e676362946d8b1d48871a3ac0b200b/do-beurteilungswesen-data.pdf).

[13] Siehe Koalitionsvertrag, Zeilen 1854-1856 zu Durchlässigkeit und Rotation.

[14] BVerwG, Urteil vom 24.11.2005, 2 C 34.04, Leitsatz Nr. 1 und Rn. 20.

[15] BVerwG, Beschluss vom 07.03.2017, 2 B 25.16, Leitsatz Nr. 3.

[16] BVerwG, Beschluss vom 07.03.2017, 2 B 25.16, Rn. 8.

hergestellt werden kann. Damit ist eine Gleichbehandlung von Frauen und Männern grundsätzlich in Frage gestellt.

II. Fehlen einer sachlichen Grundlage für Gendergerechtigkeit

Soweit die meist unbekannten Beurteilungsrichtlinien der Bundesverwaltung überhaupt verallgemeinernd untersucht werden können, spricht einiges dafür, dass sie aus verschiedenen Gründen männliche Beamt*innen bevorzugen. Dies ist sowohl am Beurteilungsverfahren als auch an den Beurteilungskriterien zu erkennen.

1. Mängel im Beurteilungsverfahren

Zur Herstellung eines einheitlichen Beurteilungsmaßstabs als Voraussetzung für Geschlechtergerechtigkeit sehen die Beurteilungsrichtlinien zwar regelmäßig sogenannte Maßstäbe- oder/und Beurteilerkonferenzen vor. Doch diese Konferenzen finden in der Bundesverwaltung meist[17] vor der Durchführung eines Regelbeurteilungsdurchgangs statt und bestehen dann in einer rein abstrakten Einführung in die Anwendung der Beurteilungskriterien. Im besten Fall kann die Beurteilerkonferenz wegen der Vielzahl der zu Beurteilenden im Zweifel nicht mehr leisten als die Vergabe von Höchstbewertungen unter einen Erklärungsvorbehalt zu stellen.

Dass die Beurteilungskriterien dann auch tatsächlich von den Beurteilenden einheitlich angewandt werden, soll durch Vorgesetzte der Beurteilenden gewährleistet werden. So ist vorgesehen, dass Vorgesetzte der Beurteilenden oder der Verfassenden von Beurteilungsbeiträgen die dienstliche Beurteilung ändern oder bestätigen, um einen einheitlichen Beurteilungsmaßstab herzustellen. Für die Personen in den Hierarchiestufen werden dabei je nach Sprachgebrauch der Beurteilungsrichtlinie unterschiedliche Bezeichnungen verwendet: Erst-, Zwischen- und Endbeurteiler*innen oder Erst-, Zweit- und Drittbeurteiler*innen oder Berichterstatter*in, Mitbeurteiler*in und Beurteiler*in. Unabhängig von den verwendeten Bezeichnungen geht dieses Verfahren von der unrealistischen Annahme aus, dass hierarchisch höherstehenden Personen Befähigung und Leistungen aller ihnen unterstehenden Personen gleichermaßen beurteilen können. Dabei kann man bezweifeln, ob sie die Befähigung und Leistungen dieser ihnen oftmals kaum bekannte Personen angemessen beurteilen sowie mit den anderen derselben Vergütungsgruppe vergleichen können, um einen einheitlichen Beurteilungsstandard herbeizuführen. Auch dieses System kann letztlich nicht mehr leisten als Ausschläge im Querschnitt der Beurteilungen nach oben bei besonders großzügigen Beurteiler*innen oder nach unten bei besonders strengen Beurteiler*innen einzufangen.

Insbesondere bei den Endbeurteilenden ist darüber hinaus auch fraglich, ob sie überhaupt fachlich qualifiziert sind, Befähigung und Leistungen von Verwaltungsbeamt*innen zu beurteilen, soweit sie nämlich selbst ihre Funktion nicht auf Grund eines verpflichtenden Auswahlverfahrens, etwa nach § 8 BBG, erlangt haben, sondern als politische Beamt*innen eingesetzt wurden. In ihrer Funktion als (End-)Beurteilende können politische Beamt*innen mittels ihrer Beurteilungspraxis bewirken, dass eine konkrete politische Linie oder wenigstens eine gewisse politische Wendigkeit zur Voraussetzung für berufliches Vorwärtskommen werden. Eine solche Beurteilungspraxis stellt das Neutralitätsgebot als hergebrachten Grundsatz des Berufsbeamtentums in Frage. Nach ständiger Rechtsprechung des

[17] Anderes regeln zum Beispiel die Beurteilungsrichtlinien in Bayern und Baden-Württemberg, die auch öffentlich zugänglich sind (s. Fußnoten 44 und 45), und mindestens eine öffentlich nicht zugängliche Beurteilungsrichtlinie eines Bundesministeriums.

Bundesverfassungsgerichts hat das Berufsbeamtentum die „ihm vom Grundgesetz zugewiesene Aufgabe, im politischen Kräftespiel eine stabile, gesetzestreue Verwaltung zu sichern“.[18]

Die Rolle politischer Beamt*innen im Beurteilungswesen wirkt sich auch konkret auf die Aufstiegschancen von Frauen aus. So überrascht nicht, was ZEIT online als die „Hans-Bremse“ beschrieben hat. Nach der Auswertung von ZEIT online im Jahr 2019 gab es in der Geschichte der Bundesrepublik Deutschland in beamteter Staatssekretärsfunktion mehr Personen mit dem Vornamen Hans als es Frauen gab. Aus dem deutlichen Übergewicht von Männern in Staatssekretärsfunktion leiten die Autor*innen des Artikels die Unterrepräsentation von Frauen in Führungspositionen der Verwaltung ab. Sie zeigen auf, dass im Laufe der Geschichte der Bundesrepublik Deutschland Positionen von beamteten Staatssekretär*innen bis zum Jahr 2009 nur in Ausnahmefällen mit Frauen besetzt wurden und dieser Frauenanteil seither auch die Schwelle von 25 % der zu besetzenden Stellen nicht überschritten hat.

Wie sich eine solche Besetzungspolitik auf die Besetzung von Führungspositionen auswirkt, legt ein Blick auf das Ressort des Bundesfinanzministeriums in den 20 Jahren von 1999 bis 2018 nahe. Im betrachteten Zeitraum waren in dem Ressort, in dem ja angeblich der Geist der schwäbischen Hausfrau herrscht, nur Männer als Staatssekretäre tätig und der Anteil der Frauen an den Abteilungsleitungen lag im Schnitt bei einer von durchschnittlich neun Abteilungen. Von allen zu vergebenden Leitungsfunktionen im Bundesfinanzministerium wurden 135 mit Männern und 18 mit Frauen besetzt.[19] Zu solcher extremen Überzahl von Männern wird traditionell[20] eingewandt, es fehle eben noch an qualifiziertem weiblichem Nachwuchs. Um dieses Argument auf seine Stichhaltigkeit zu prüfen, erhebt das Statistische Bundesamt im Gleichstellungindex auch Zahlen zum Anteil von Frauen an der gesamten Belegschaft von Behörden und speziell im höheren Dienst. Diese Zahlen zeigen deutlich auf, dass das Reservoir an Frauen im höheren Dienst noch lange nicht für Führungspositionen ausgeschöpft ist.[21]

Ob man eine ähnliche Gefahr der Einflussnahme zum Nachteil von Frauen, wie sie ZEIT online bei Staatssekretärsfunktionen beobachtet hat, auch bei anderen, politisch besetzten Führungspositionen, wie etwa die Leitungen des Kommunikationsbereichs eines Ministeriums, annehmen kann, ist noch nicht untersucht worden. Abgesehen von der auch in diesen Fällen anzunehmenden Gefahr für ein politisch neutrales, allein Recht und Gesetz verpflichtetes Beamtentum, ist hier aus der Perspektive der Geschlechtergerechtigkeit von Interesse, ob sich das Beurteilungsverfahren allgemein zum Nachteil von Frauen auswirken kann.

So klar strukturiert das in der Bundesverwaltung praktizierte Beurteilungsverfahren wirkt, fehlt ihm doch auffällig ein Element, das Frauen in Hierarchien vermissen: der Diskurs.[22] In den meisten der hier aus verschiedenen Geschäftsbereichen der Bundesverwaltung bekannten Beurteilungsrichtlinien ist zwar vorgesehen, dass die erstbeurteilende Person bzw. die Person, die einen Beurteilungsbeitrag zu

---

[18] BVerfGE 121, 205ff. (221) m.w.N.; zu der Frage, wie die Rolle politischer Beamter im Verhältnis zum verfassungsrechtlichen Neutralitätsgebot der Verwaltung steht, im Einzelnen *Marianne Czisnik*, Die verfassungsrechtliche Stellung der politischen Beamten, Die öffentliche Verwaltung [DÖV] 2020, S. 603-612 (604).

[19] https://www.zeit.de/politik/deutschland/2018-09/gleichberechtigung-frauen-diskriminierung-fuehrungspositionen-ministerien.

[20] So auch in den Kommentaren im Internet zum zitierten Artikel.

[21] https://www.destatis.de/DE/Themen/Staat/Oeffentlicher-Dienst/Publikationen/Downloads-Oeffentlicher-Dienst/gleichstellungsindex-5799901247004-1_2024265.pdf?__blob=publicationFile&v=3, S. 9; anschaulich, wenn auch etwas veraltet sind auch die Diagramme bei https://www.zeit.de/politik/deutschland/2018-09/gleichberechtigung-frauen-diskriminierung-fuehrungspositionen-ministerien; siehe auch die Problemdarstellung in der Einleitung.

[22] *Doris Bischof-Köhler*, „Geschlechtstypische Besonderheiten im Konkurrenzverhalten: Evolutionäre Grundlagen und entwicklungspsychologische Fakten“; in: Gertraude Krell / Margit Osterloh (Hrsg.), Personalpolitik aus der Sicht von Frauen. Was kann die Personalforschung von der Frauenforschung lernen? (Sonderband der Zeitschrift für Personalforschung, 1992), S. 251-281, unterscheidet in (von Männern gebildete) Dominanzhierarchien und (von Frauen bevorzugte) Ranghierarchien, siehe S. 261, 268-270.

entwerfen hat, vor der Beurteilungserstellung ein Gespräch mit der zu beurteilenden Person führt.[23] In diesem Gespräch geht es allerdings nur um die darzustellenden Aufgaben und gegebenenfalls besonderen Herausforderungen im Beurteilungszeitraum. Über die für die dienstliche Beurteilung zu berücksichtigenden Leistungen braucht in diesen Fällen weder in dem Gespräch vor der Erstellung der dienstlichen Beurteilung noch im Laufe des zu betrachtenden, vorangegangenen Beurteilungszeitraums gesprochen zu werden. Dieses Nichtvorsehen von Diskurs kann für die zu beurteilende Person zu einer Überraschung hinsichtlich des Ergebnisses der dienstlichen Beurteilung führen. Sie kann außerdem mögliche Leistungsschwächen mangels Diskurses auch nicht abstellen. Diese fehlende Auseinandersetzung im Gespräch ist also nicht nur befremdlich für Frauen, sondern widerspricht auch dem öffentlichen Interesse an Leistungserbringung nach Art. 33 Abs. 2 GG.

Auch zwischen den beurteilenden Personen der verschiedenen Hierarchiestufen fehlt es an der Verpflichtung zum Diskurs. Die Beurteilungsrichtlinien verlangen keinen Austausch zur Plausibilisierung der dienstlichen Beurteilung. Es ist daher fragwürdig, ob die verschiedenen mit der dienstlichen Beurteilung derselben Person betrauten Personen jeweils von einer hinreichenden und angemessen überprüften Tatsachengrundlage ausgehen.[24] Das übliche Beurteilungssystem unterstellt offenbar überlegene Urteilsfähigkeit auf der höheren Hierarchiestufe. Mit dieser Vereinzelung der jeweils an einer dienstlichen Beurteilung mitwirkenden Person, die sich für ihre Wertungen nicht erklären muss, ist das Beurteilungswesen offen für sachfremde Erwägungen und sogar für bewusste Manipulation. Diese Gefahren stellen die Sachgerechtigkeit der dienstlichen Beurteilungen an sich in Frage.

Zusammenfassend kann man das Beurteilungsverfahren als paternalistisch beschreiben – es ist intransparent und in Teilen bevormundend. Diese Kombination von Charakteristika lässt sich aus den überkommenen Vorstellungen des besonderen Gewaltverhältnisses erklären, die unserer Grundrechtsordnung nicht entsprechen. Die Lehre vom besonderen Gewaltverhältnis war schon zur Zeit ihrer Entstehung Ende des 19. Jahrhunderts eher eine soziologische Beobachtung als eine aus Rechtsgrundsätzen ableitbare Theorie. Sie wurde üblicherweise auf Strafgefangene, Schüler*innen und Beamt*innen angewandt, die nach dieser Lehre unter Verlust gewisser Freiheitsrechte in ein Gehorsamsverhältnis eingebunden sind. Nur die Beamt*innen traten freiwillig in das sie betreffende Gewaltverhältnis ein. Der Staatsrechtler Paul Laband, der den Begriff des besonderen Gewaltverhältnisses prägte, verglich den Eintritt in das Beamtenverhältnis mit einer Unterwerfung unter „väterliche Gewalt“. In diesem Verhältnis unterwarf sich der (damals noch rein männlich zu denkende) Beamte einer „besonderen Gehorsams-, Treue- und Dienstpflicht“, um „Schutz“ (wovor auch immer) und Diensteinkommen gewährt zu bekommen.[25] Die Vorstellung von einem besonderen Gewaltverhältnis erfolgreich in Frage zu stellen, bedurfte es eines Strafgefangenen, der sich aus dem für ihn sehr unausgewogenen Gewaltverhältnis zu befreien suchte, indem er das Grundrecht auf Briefgeheimnis für sich in Anspruch nahm. Im Strafgefangenen-Urteil von 1972 urteilte das Bundesverfassungsgericht, dass auch Personen in Sonderrechtsverhältnissen die Grundrechte prinzipiell zustehen. Die Grundrechte dürfen in solchen Rechtsverhältnissen nur durch Gesetz eingeschränkt werden und zwar auch nur so weit, wie der Zweck des betreffenden Sonderrechtsverhältnisses dieses erfordert.[26] Dieses Urteil führte zur Verabschiedung des Strafvollzugsgesetzes im Jahr 1977 und wirkte sich auch auf die Rechtsstellung von Heimkindern und Schüler*innen aus. Im Bereich des Beamtenrechts bediente sich allerdings noch 16 Jahre nach dem Strafgefangenenurteil des Bundesverfassungsgerichts der einflussreiche Beamtenrechtlicher Walter

[23] Unterschiedliche Bezeichnungen aus den verschiedenen Beurteilungsrichtlinien (wie etwa Berichterstatter*in – Mitbeurteiler*in – Beurteiler*in) werden hier der sprachlichen Einfachheit wegen vernachlässigt.
[24] BVerwG 2 A 10.17 / 2 C 21.16, Leitsatz Nr. 1.
[25] Paul Laband, Das Staatsrecht des Deutschen Reiches (3 Bd., 1876), Bd. 1, Seite 387.
[26] BVerfG 2 BvR 41/71, Beschluss vom 14.03.1972.

Wiese[27] des Vokabulars des besonderen Gewaltverhältnisses, wenn er etwa schrieb: „Niemand ist gezwungen, ein derartiges Amt zu übernehmen. [...] Hat ein Bürger von Freiheit und Recht in diesem Sinne Gebrauch gemacht, dann hat er in gleichem Maße seine grundrechtliche Ausgangsposition verbraucht; er darf nunmehr seine ursprünglichen Freiheiten nicht mehr gegen den neuen Status ausspielen“. Zur Veranschaulichung zitierte er eine sehr männliche Sichtweise auf die Rechtsposition einer verbeamteten Person: „Man kann nicht Ehemann unter Vorbehalt der Rechte des Junggesellen sein“.[28]

Ein Beispiel, dass eine verbeamtete Person in der Ausübung ihres Dienstes im Vergleich zu anderen Staatsbürger*innen weniger geschützt ist, findet sich auch noch im Strafgesetzbuch, wo § 193 StGB regelt: „Vorhaltungen und Rügen der Vorgesetzten gegen ihre Untergebenen, dienstliche Anzeigen oder Urteile vonseiten eines Beamten und ähnliche Fälle sind nur insofern strafbar, als das Vorhandensein einer Beleidigung aus der Form der Äußerung oder aus den Umständen, unter welchen sie geschah, hervorgeht.“ Wer sich verbeamten lässt, begibt sich also in gewissem Umfang eines Rechtsschutzes vor strafbarer Handlung, der außerhalb dieses besonderen Raums anderen Bürgern gewährt wird.[29] Selbst in dem unbestreitbaren Bereich der Organisationshoheit des Dienstherrn lässt die Wortwahl eine ungehemmte Verfügungsmacht über Beamt*innen vermuten, etwa wenn nicht aktiv von Aufgabenwahrnehmung, sondern passiv, wie bei einem Gegenstand, von „Verwendung“ die Rede ist.[30]

Manche hergebrachten Vorstellungen vom Beamtentum wirken sich damit auf das Beurteilungsverfahren in einer Weise aus, die verbeamtete Personen in eine passive oder bestenfalls reagierende, jedenfalls einem fest gefügten System unterworfene Rolle zwingen, ohne dass ihre Gestaltungsfähigkeit gefragt wäre. Diese Vorgehensweise, die einen Austausch über Erkenntnisgrundlagen und Erwartungshaltungen zur Leistungserbringung nicht vorsieht und einen für Frauen vorzugswürdigen Diskurs nicht ermöglicht, wirkt sich auch auf die Anwendung der Beurteilungskriterien aus.

2. Mängel bei den Beurteilungskriterien

Unabhängig von den formellen Schwierigkeiten bei der dienstlichen Beurteilung sind auch die konkreten Beurteilungskriterien zum Teil ungeeignet, Geschlechtergerechtigkeit zu gewährleisten. So hat die nach Art. 23 Abs. 3 Führungspositionengesetz (FüPoG) vorgesehene Evaluation ganz allgemein ergeben, dass je größer die Zahl der Beurteilungskriterien, umso höher ist die Wahrscheinlichkeit für eine Anwendung zum Nachteil von Frauen. Hintergrund ist, dass eine sehr kleinteilige Betrachtung von Leistungskriterien leichter zu einem bei der Beförderungsentscheidung maßgeblichen Vorsprung der Wunschbewerberin oder des Wunschbewerbers bei gleicher Endnote mit Konkurrent*innen führt. Die

[27] In seinem „Handbuch des Öffentlichen Dienstes“ wird er beschrieben als: „Ministerialrat im Bundesinnenministerium [mit] zahlreiche[n] Verwendungen im Landes- und Bundesdienst, u.a. Mitglied der Arbeitsgruppe ‚Reform des Öffentlichen Dienstrechts‘ im Bundesinnenministerium“.

[28] Walter Wiese [Hrsg. und Autor], Handbuch des Öffentlichen Dienstes, Band II Teil 1 Beamtenrecht (3. vermehrte Auflage, Köln, Berlin, Bonn, München: Carl Heymanns Verlag KG, 1988), Seite 24; die Passage bezieht sich auf das Nebentätigkeitsrecht, wobei dem Autor die Rechtsprechung des Bundesverfassungsgerichts bekannt zu sein scheint, wenn er meint sich erklären zu müssen: „Von einer Grundrechtseinschränkung kann in derartigen Fällen nicht die Rede sein.“ Art. 12 GG, Berufsfreiheit, scheint ihm nicht beachtenswert.

[29] Der systematische Ansatz, bei dem das Recht nicht oder weniger schützt, wenn die Tat in einem Raum begangen wird, der von der Öffentlichkeit abgegrenzt ist und in dem ein Machtgefälle besteht, lässt sich mit der mit der Straffreiheit der Vergewaltigung in der Ehe vergleichen, gegen die der djb sich so erfolgreich eingesetzt hat, auch wenn hier der Eingriff in die Rechtsposition eines anderen deutlich weniger erheblich ist.

[30] Siehe etwa in Fußnote 27.

Bundesregierung hat dementsprechend im Jahr 2020 für sich anerkannt zu prüfen, wie die Ausdifferenzierung der Beurteilungskriterien begrenzt werden kann.[31]

Die Diversifizierung von Leistungskriterien führt tendenziell zu einer Fixierung auf einzelne Aspekte der Leistungserbringung, die oft von zweitrangiger Bedeutung sind. Die wenigsten Beurteilungsverordnungen gehen von der Gewichtung von Kriterien aus. Das führt dazu, dass ein Kriterium wie das mündliche Ausdrucksvermögen gleichrangig etwa mit Fachkenntnissen bewertet wird. Andere Kriterien wiederum geben Raum, immer wieder dieselbe Leistung zu bewerten; typischerweise geschieht das bei Kriterien wie Entschlusskraft, Arbeitsplanung und Belastbarkeit, sich sich anhand der reinen Arbeitsmenge bemessen lassen. Damit wird ein Aspekt der Leistung, der von Beamt*innen ohne Care-Aufgaben leichter zu erfüllen ist, auf gleich drei Kriterien verteilt und bekommt damit ein erhöhtes Gewicht. In einem solchen System gerät im Übrigen eine überblicksartige Wertung – und damit Würdigung – aus dem Blick, ob die beurteilte Person selbständig Verantwortung übernimmt, etwa durch strategisches Vorgehen und das Berücksichtigen von Zusammenhängen und Folgen des Verwaltungshandelns.

Die Konzentration auf kontrollierbare Einzelaspekte, die viele Beurteilungsrichtlinien auszeichnet, spiegelt das Interesse an einer paternalistischen Kontrolle wider. Die beurteilte Person wird als überprüfbar und reaktiv, nicht als gestaltend wahrgenommen und entsprechend auf fehlerfreie Ausführung von Weisungen geprüft, anstatt in der Gestaltung ihres Verantwortungsbereichs wahrgenommen und gewürdigt zu werden. Dass die beurteilte Person Zusammenhänge berücksichtigt und Folgebetrachtungen anstellt – kurz, Verantwortung übernimmt – findet keine oder allenfalls beiläufige Beachtung. Zwischenmenschliche Teamfähigkeit und inhaltliche Priorisierungsfähigkeit können mit dem Perfektionsanspruch, den andere Kriterien (bis hin zum sprachlichen Ausdruck) voraussetzen, nicht abgebildet werden. Im Ergebnis verhindern die Wertungen von Beurteilungssystemen innovative Lösungsansätze, die sich die neue Bundesregierung gerade wünscht, und verfestigen Vorstellungen von Herrschaft und Hierarchie, in denen Frauen sich tendenziell weniger gut zurechtfinden.

Ein Beispiel für eine paternalistische Sichtweise ist die Bedeutung, die manche Beurteilungsregelungen der „Delegation“ zumessen. Abstrakt gesehen geht es bei diesem Leistungskriterium um die Festlegung von Zuständigkeiten. Man könnte also theoretisch auch von der „Festlegung von klaren Zuständigkeiten im Verantwortungsbereich“ sprechen. Solche Festlegungen von Zuständigkeiten sollten Ergebnis eines möglicherweise umfangreichen Entscheidungsprozesses sein, bei dem sich die vorgesetzte Person mit den Auswirkungen der aktuellen Aufgabenverteilung auseinandersetzt und an dem die Mitarbeitenden durchaus beteiligt sein können, etwa indem sie den Arbeitsanfall in ihrem jeweiligen Zuständigkeitsbereich beschreiben. In der Praxis wird „Delegation“ dagegen meist nicht so komplex in ihren Wechselwirkungen betrachtet, sondern als Machtinstrument zur Übertragung einzelner Aufgaben benutzt. Im Ergebnis solcher Praxis fühlen sich Mitarbeitende oftmals unvorhersehbaren Aufgabenübertragungen ausgeliefert und durch eine Vielzahl von zum Teil überraschenden Delegationen verunsichert. Jedenfalls lässt sich feststellen, dass je mehr Delegierungsentscheidungen eine Führungskraft trifft, umso mehr dafür spricht, dass sie keine klaren Zuständigkeiten festgelegt hat, keine transparente Strategie verfolgt und willkürlich vorgeht. Häufiges Delegieren kann damit durchaus ein Zeichen für schlechte Leistungen einer Führungskraft und ein äußerst fragwürdiges Beurteilungskriterium.

[31] Stellungnahme der Bundesregierung zu Evaluationsgutachten über die Wirksamkeit des Gesetzes für die gleichberechtigte Teilhabe von Frauen und Männern an Führungspositionen in der Privatwirtschaft und im öffentlichen Dienst sowie über den Erfüllungsaufwand, BT-Drucksache 19/24615, S. 14.

Sogar solche Regelungen sind nicht per se gleichstellungsfreundlich, nach denen spezifische, durch Familien- oder Pflegeaufgaben und in der sozialen Arbeit erworbene Erfahrungen und Fähigkeiten bei der Bewertung der Eignung und Befähigung zu berücksichtigen sind, soweit diese Fähigkeiten für die Ausübung der jeweiligen Tätigkeit von Bedeutung sind.[32] Solche Regelungen mögen vordergründig gleichstellungsfreundlich sein, sind aber in der Praxis bislang kaum anwendbar. Der Dienstherr darf nach familiären Umständen nämlich nicht fragen, die Beurteilerin oder der Beurteiler erfahren von

besonderen Leistungen oder Lasten also nur zufällig oder wenn die Bewerber*innen ihre durch Pflege- und Familienaufgaben erworbene besondere Qualifikation für den konkreten Arbeitsplatz selbst darstellen, was sich schlecht objektivieren lässt. Zugleich können nahezu beliebig Fähigkeiten zugeschrieben werden. Das Evaluationsgutachten über die Wirksamkeit des Gesetzes für die gleichberechtigte Teilhabe von Frauen und Männern an Führungspositionen in der Privatwirtschaft und im öffentlichen Dienst - FüPoG I -⁻ weist auf den problematischen Gehalt derartiger Regelungen ausführlich hin und nennt als Beispiel einen „stets berufstätigen Vater von drei Kindern“, dem „ein besonderes Organisationstalent zugeschrieben werden [kann], mit der Folge, dass er sich gegen eine weibliche Konkurrentin durchsetzt“. Schließlich ist schwer zu greifen, bei welchen Tätigkeiten die Fähigkeiten von Bedeutung sein können. Um gleichwohl beurteilungsrelevante Erfahrungen und Qualifikationen aus Care-Arbeit oder Ehrenamt[33] diskriminierungsfrei berücksichtigen zu können, sind Kriterien und Verfahren erforderlich. An diesen fehlt es, soweit ersichtlich, bislang.

Selbst wenn die Befähigungs- und Leistungskriterien zahlenmäßig überschaubar und sinnvoll ausgewählt sind, stecken noch einige Gefahren für das geschlechtergerechte Beurteilen in der Anwendung der Kriterien. Ein klassischer Beurteilungsfehler liegt im sogenannten Ähnlichkeitseffekt. In den so bezeichneten Fällen „werden Personen, die der Beurteilerin oder dem Beurteiler ähnlich sind, von ihr oder ihm besser eingeschätzt“.[34] Die Vorstellung, wie eine leistungsstarke Person zu sein hat, wird auf diese Weise aus den tatsächlichen, meist männlich geprägten Gegebenheiten abgeleitet und nicht – wie es sein sollte – aus ergebnisorientierten Erwägungen, die eine Diversität an Arbeitsstilen zulassen.

Andere Effekte, die einer möglichst objektiven Herangehensweise an die dienstliche Beurteilung einer Person entgegenstehen, sind zum Beispiel der Hierarchie-Effekt, nach dem Mitarbeitende in höheren Funktionsgruppen per se besser eingeschätzt werden, der Exzellenz-Effekt, nach dem Männer eher als exzellent eingestuft werden als Frauen, der für Frauen offensichtlich besonders relevante Teilzeit-Effekt, der dazu führt, dass Mitarbeitende in Teilzeit als weniger leistungsfähig eingeschätzt werden, und der Eltern-Effekt, nach dem Väter besser eingeschätzt werden als Mütter und Frauen ohne Kinder besser als Mütter. *Frey* hat diese und weitere Effekte den üblichen Leistungsmerkmalen aus Beurteilungsrichtlinien zugeordnet und damit für die Auswirkungen eines gender-bias sensibilisiert.[35]

Ähnlich abträglich für ein geschlechtergerechtes Beurteilungswesen wie die skizzierten abstrakten Effekte sind konkrete, männlich attribuierte Eigenschaften bzw. Erwartungen. Hierzu zählen anerkanntermaßen Durchsetzungsfähigkeit und Engagement. Auch Identifikation mit der Aufgabe und „termingerechtes bzw. sorgfältiges Arbeiten“ assoziieren diffuse Vorstellungen von umfassender

---

[32] Zur Kritik vgl. djb, Stellungnahme 24-28 vom 30.07.2024 zum Entwurf eines Gesetzes zur Verwirklichung der Gleichberechtigung (Neufassung des Niedersächsischen Gleichberechtigungsgesetzes), S. 6 f., abrufbar unter https://www.djb.de/presse/stellungnahmen/detail/st24-28.

[33] vgl. zur Berücksichtigung der „besonderen persönlichen Lage von Bewerbern“ bei Beurteilungen EuGH, Urt. v. 28.03.2000, Rs. C-158/97 (Badeck), Rn. 31 bis 27; Urt. v. 06.07.2000, Rs. C-407/98 (Abrahamsson), Rn. 61.

[34] *Regina Frey*, Die dienstliche Beurteilung. Fair beurteilen, gleichstellungsorientiert handeln. Leitfaden für Beurteilerinnen und Beurteiler sowie Personalverantwortliche der Landesverwaltung Sachsen-Anhalts² (Sachsen-Anhalt, Ministerium für Justiz und Gleichstellung, 2015), S. 20 m.w.N.

[35] *Frey* (2015), S. 18-21 (zu den verschiedenen Effekten) und 23 (zu den Auswirkungen der Verzerrungseffekte auf die Bewertung von Leistungsmerkmalen).

zeitlicher Verfügbarkeit und leistungsbezogenen Präsentismus, die eher Männern zugeschrieben werden.[36] *Lorse* schlägt daher als „weitere[n] Baustein auf dem Weg zur geschlechtsindifferenten dienstlichen Beurteilung“ vor, die Leistung verstärkt unter dem Aspekt eines eindeutigen Aufgaben- und Ergebnisbezugs („output orientiert“) zu bewerten“.[37] Dieser Aspekt sollte dann aber auch nur einmal bewertet werden und nicht versteckt in unterschiedlichen Kriterien.

Unabhängig davon, von welchem Anwendungsfehler die größte Gefahr für eine geschlechtergerechte dienstliche Beurteilung ausgeht, steht der Anwendung von Beurteilungskriterien oftmals ganz grundsätzlich das Interesse der vorgesetzten Person entgegen, allenfalls Mitarbeitenden eine Laufbahn im öffentlichen Dienst zu ermöglichen, die höchstens so leistungsstark wie sie selbst sind und nach Möglichkeit niemanden, der auf eine andere Weise erfolgreich ist. So verhindert die beurteilende Person, bewusst oder unterbewusst, dass sich ihre eigenen Schwächen zeigen. Diese Tendenz beeinträchtigt die Leistungsqualität in der Verwaltung. Eine solche Entwicklung verstärkt die bereits bestehenden Nachteile für Frauen.

C. Lösungen

Die festgestellten Mängel tragen ihre Behebung oftmals schon in sich, nach dem Motto „Gefahr erkannt, Gefahr gebannt“. Hinzu kommt, dass angesichts der Vielfalt an Beurteilungssystemen positive Ansätze für ein geschlechtergerechtes Beurteilungswesen oftmals schon in einzelnen Bundesressorts oder Landesverwaltungen praktiziert werden. Lösungen sind also realistisch erreichbar.

I. Vereinheitlichung der Beurteilungsregelungen

Die erste Hürde auf dem Weg zum geschlechtergerechten Beurteilen dürfte darin bestehen, statt der bisherigen Intransparenz und Zersplitterung einer Vielzahl von Beurteilungssystemen auf Bundesebene zu einer Vereinheitlichung zu kommen. Idealerweise sollte es nur eine Beurteilungsrichtlinie für die gesamte Bundesverwaltung (bzw. jeweilige Landesverwaltung) geben. Dies ist auch Ziel der neuen Bundesregierung. In ihrem Koalitionsvertrags „Verantwortung für Deutschland“ setzt sich die neue Bundesregierung das Ziel „einheitliche Beurteilungsstandards für die Bundesverwaltung“ zu schaffen.[38] Dies wäre auch mit dem Vorteil verbunden, dass nicht in einer Vielzahl von Beurteilungsrichtlinien, sondern nur in einer Beurteilungsrichtlinie Anpassungen vorgenommen werden müssen, wenn etwa die Rechtsprechung einen Änderungsbedarf ergibt.

Schwieriger als die Bestimmung des Ziels gestaltet sich die Umsetzung, zumal die Zuständigkeit fraglich ist. Auf Bundesebene könnte das neu geschaffene Ministerium für Digitalisierung und Staatsmodernisierung (BMDS) in Frage kommen, zumal die „Initiative für einen handlungsfähigen Staat“, auf die sich der Koalitionsvertrag bezieht,[39] vorschlägt, dass Querschnittsaufgaben ressortübergreifend gebündelt werden.[40] Das BMDS wird aber offenbar nicht die mit der nötigen Kompetenz ausgestattete Dienstrechtsabteilung aus dem Bundesministerium des Innern übertragen bekommen. Unabhängig davon, welches Ressort mit der Schaffung „einheitliche[r] Beurteilungsstandards“ betraut werden soll, geht wohl auch die Bundesregierung in ihrem

[36] *Lorse* (2017), S. 459.
[37] *Lorse* (2017), S. 459.
[38] koav_2025.pdf, Zeile 1833.
[39] koav_2025.pdf, Zeile 1792.
[40] Abschlussbericht der „Initiative für einen handlungsfähigen Staat“, S. 52 (https://www.ghst.de/fileadmin/images/01_Bilddatenbank_Website/Demokratie_staerken/Initiative_f%C3%BCr_einen_handlungsf%C3%A4higen_Staat/Abschlussbericht/Abschlussbericht_neu.pdf).

Koalitionsvertrag davon aus, dass „Gebündelte Service-Einheiten statt Doppelstrukturen“ zur Erreichung der angestrebten „Modernisierung des öffentlichen Dienstrechts“[41] nötig sind. Eine zentrale Zuständigkeit lässt sich idealerweise in der Gemeinsamen Geschäftsordnung der Bundesregierung (GGO) regeln. Eine solche zentrale Zuständigkeit sollte umgehend geschaffen werden, um die Voraussetzung für die Vereinheitlichung des Beurteilungswesens und damit der Gleichbehandlung von Frauen und Männern in der Bundesverwaltung zu schaffen. Gleiches ist für die Länder zu fordern.

Fachliche Gegenargumente gegen eine Vereinheitlichung des Beurteilungswesens vermögen nicht zu überzeugen. Wenn die Zersplitterung der Beurteilungsregelungen in der Bundesverwaltung sich auch mit dem Ressortprinzip nach Art 65 S. 2 GG erklären lässt, verlangt das Ressortprinzip doch keinesfalls, dass es keine ressortübergreifenden Standards geben darf. Schon in der aktuellen gemeinsamen Geschäftsordnung der Bundesregierung (GGO) ist in § 19 Abs. 1 S. 1 festgelegt: „In Angelegenheiten, die die Geschäftsbereiche mehrerer Bundesministerien berühren, arbeiten diese zusammen, um die Einheitlichkeit der Maßnahmen und Erklärungen der Bundesregierung zu gewährleisten.“ Außerdem dürften für ein einheitliches Beurteilungssystem (mindestens) in der Bundesverwaltung das Gleichheitsgebot nach Art. 3 GG und das Leistungsprinzip nach Art. 33 Abs. 2 GG sprechen. Es ist also völlig legitim, wenn der Koalitionsvertrag zum Rückbau von Bürokratie eine ressortübergreifende Zusammenarbeit anstrebt.[42]

Auch stehen möglicherweise unterschiedliche Anforderungen in Geschäftsbereichen einer Vereinheitlichung nicht entgegen, wenn diese Vereinheitlichung gleichzeitig mit einer Reduzierung der Leistungskriterien verbunden wird, von der sogleich unter II.2 noch zu handeln sein wird.

II. Sachliche Beurteilungsgrundlage zur Schaffung von Gendergerechtigkeit

1. Sachlichkeit im Beurteilungsverfahren

Ansätze zur Verbesserung von Beurteilungsverfahren müssen die festgestellten Mängel in einigen der praktizierten Beurteilungssysteme beseitigen helfen, also den politischen Einfluss zurückdrängen und mehr Diskurs zu Leistungserwartungen und Erkenntnisgrundlagen ermöglichen.

Die strukturelle Abhängigkeit von politischer Einflussnahme, erinnert sei an die „Hans-Bremse“, muss verhindert werden, um Geschlechtergerechtigkeit und die Neutralität des Beamtentums zu gewährleisten. Die dienstliche Beurteilung darf nicht als Belohnung für Unterwerfung unter politischen Willen interpretiert werden können. Hierfür ist es erforderlich, Personen, die ihr Amt ohne ein Auswahlverfahren nach § 8 BBG erlangt haben, von Personalentscheidungen allgemein und von der Rolle einer beurteilenden Person im Besonderen auszuschließen. Ein solches Vorgehen wird zum Beispiel in Großbritannien praktiziert. Dort steht als oberste beurteilende Person an der Spitze des nationalen Verwaltungsapparats ein*e von Regierungswechseln unabhängige*r head of civil service.[43]

Ein einheitlicher Beurteilungsstandard kann aber auch erreicht werden, ohne dass eine Person an der Spitze einer Behörde ihn herstellt. So wird in den transparent im Internet dargestellten

[41] koav_2025.pdf, Zeilen 1828 bis 1833, 1846 bis 1856.
[42] https://www.koalitionsvertrag2025.de/sites/www.koalitionsvertrag2025.de/files/koav_2025.pdf, Zeilen 1822-1827.
[43] *Czisnik* (2020), S. 609.

Beurteilungssystemen in Bayern[44] und Baden-Württemberg[45] und mindestens einem, nur Eingeweihten bekannten, Bundesressort eine Kommission als neutrales Gremium betraut, um nach der Erstellung der Beurteilungsentwürfe die Einhaltung des Vergleichsmaßstabs zu gewährleisten. Der in einem solchen Gremium stattfindende Diskurs macht es wahrscheinlicher, dass sich die Entscheidungsträger um eine sachliche Erkenntnisgrundlage bemühen, dass Leistungserwartungen geklärt werden und dass Interessen einzelner Vorgesetzter ausgeglichen werden. In Baden-Württemberg schreibt die entsprechende Verwaltungsvorschrift sogar eine paritätische Besetzung der Beurteilungskommission vor.[46] Zur Umsetzung des diskriminierungsfreien Beurteilens auch in einer zur Versachlichung beitragenden Beurteilungskommission, die für die Herstellung eines einheitlichen Beurteilungsmaßstabs zuständig ist, empfiehlt sich die Verpflichtung zur Teilnahme an Fortbildungsmaßnahmen zum geschlechtergerechten Beurteilen.[47]

Die Probleme der Vergleichsgruppenbildung dürften auch für kleine Behörden erleichtert werden, wenn die Vereinheitlichung des Beurteilungswesens auch sie umfasst. Es gibt schon jetzt mindestens in einem Bundesressort eine Beurteilungsrichtlinie, deren Geltungsbereich den unmittelbaren Geschäftsbereich mit umfasst. Das zeigt, dass Beurteilungsrichtlinien auch ohne weiteres verschiedene Arten von Aufgaben (ministeriell und nicht-ministeriell, vollziehend, innere Verwaltung usw.) umfassen kann.

Um dem Eindruck von Willkür im Zeitpunkt der Beurteilungserstellung entgegenzuwirken, sollte vorgesehen werden, dass die zu beurteilende Person von der für sie zuständigen erstbeurteilenden Person auf wahrgenommene Leistungsverschlechterungen hingewiesen wird. Hierzu gibt es in bestehenden Beurteilungsrichtlinien von Bundesministerien bereits allgemeine Festlegungen. So heißt es in einer Beurteilungsrichtlinie zur Verantwortung der Führungskräfte, sie sollten „im strukturierten Dialog während des Beurteilungszeitraumes Potenziale der Beschäftigten erkennen, Leistungs- und Entwicklungsziele aufzeigen, aber auch Leistungsschwächen oder Probleme aufgreifen, um den Beschäftigten Gelegenheit zu geben, diese auszugleichen oder zu beheben und sie dabei zu unterstützen“. Zu dieser Vorgabe heißt es ergänzend, ihre Vernachlässigung stelle „einen Führungsmangel dar und [sei] bei allen Arten von Leistungsbewertungen und Beurteilungen der Führungskräfte zu berücksichtigen.“ Nützlicher als eine solche allgemeine Aussage ist wahrscheinlich die von einem anderen Bundesressort gemachte konkrete Vorgabe, nach der Hälfte des Beurteilungszeitraums eine Zwischeneinschätzung abzugeben, an der sich die zu beurteilende Person orientieren kann. In einem solchen Gespräch kann die zu beurteilende Person leistungserschwerende Rahmenbedingungen thematisieren, so dass die Führungskraft dazu beitragen kann, diese zu beseitigen. Im Ergebnis des Gesprächs kann die zu beurteilende Person ihre Leistungen besser an Erwartungen anpassen. Auf diese Weise können Missverständnisse und womöglich sogar Meinungsverschiedenheiten ausgeräumt werden, so dass womöglich sogar gerichtliche Streitigkeiten vermieden werden können. Ein laufender Austausch über Leistungserwartungen und Erfüllungsvoraussetzungen für Leistungserbringung liegt im Übrigen auch im öffentlichen Interesse an der Leistungserbringung nach Art. 33 Abs. 2 GG.

---

[44] In Bayern ist die Möglichkeit, solche Kommissionen zu schaffen, nach Art. 60 Abs. 1 S. 5 Leistungslaufbahngesetz (LlbG) eröffnet, https://www.gesetze-bayern.de/Content/Document/BayLlbG-60; weiteres hierzu ist zu entnehmen: https://www.gesetze-bayern.de/Content/Document/BayVwV230264-NN4; Gebrauch gemacht hat von der Möglichkeit Beurteilungskommissionen einzusetzen z.B. das Bayerische Landwirtschaftsministerium: https://www.gesetze-bayern.de/Content/Document/BayBeurtRFLF-36.

[45] § 6 Abs. 4 BeurtVO BW: „Die Einhaltung der Richtwerte nach Absatz 2 sowie der Differenzierung nach Absatz 3 ist durch Beurteilungskommissionen bei den Endbeurteilerinnen oder Endbeurteilern sicherzustellen“ (https://www.landesrecht-bw.de/bsbw/document/jlr-BeamtBeurtVBW2014V2P1).

[46] *Lorse* (2017), S. 458, Fn. 23.

[47] *Lorse* (2017), S. 463.

Eine weitere Möglichkeit, mit Vorgaben zum Verfahren zu einer Versachlichung des Beurteilungswesens beizutragen und damit unterschiedlichen Formen des gender-bias vorzubeugen, besteht in einer Verstärkung der Rolle der zu beurteilenden Person. Werden Beamt*innen als mündige Staatsbürger wahrgenommen, in das Verfahren einbezogen und nicht nur als Empfänger der Wertungen von Vorgesetzten gesehen, können sie etwa zu den Erkenntnisgrundlagen der erbrachten Leistungen für die dienstliche Beurteilung beitragen. Ein solcher Vorschlag hat beim 73. Deutschen Juristentag (DJT), Abteilung „Justiz. Empfehlen sich Regelungen zur Sicherung der Unabhängigkeit der Justiz bei der Besetzung von Richterposten?“ eine Mehrheit gefunden. Dort wurde für Richterbeurteilungen beschlossen, was für Beamtenbeurteilungen erst recht Geltung finden kann, nämlich: „Eine Beurteilungsgrundlage können Berichte des oder der zu Beurteilenden über die eigenen Leistungen sein“. Jedenfalls sollen nach einem anderen Beschluss derselben Abteilung des 73. DJT „die tatsächlichen Erkenntnisgrundlagen einer Beurteilung [...] in der Beurteilung offengelegt werden“.[48] Durch eine solche Verfahrensweise würde zur Versachlichung beigetragen und Effekten entgegengewirkt, die für Frauen nachteilig sind. Gleichzeitig erleichtert eine klare Erkenntnisgrundlage Diskussion und Entscheidung in einer Kommission.

Ein Mittel zur Förderung ist auch die Regelung, Frauen bei gleicher Eignung vorzuziehen, wie es etwa § 8 Abs. 1 BGleiG vorschreibt: „Sind Frauen in einem Bereich unterrepräsentiert, so hat die Dienststelle sie bei gleicher Qualifikation wie ihre Mitbewerber bevorzugt zu berücksichtigen“. *Lorse* weist allerding nach, wie die Rechtsprechung Detaillierungen der Leistungsbeurteilung fordert, die schließlich den möglichen Anwendungsbereich für eine „gleiche Qualifikation“ extrem reduzieren. So hat das OVG NRW entschieden, „die Identität der Gesamtnoten [...] in den aktuellen Regelbeurteilungen des Antragstellers und der Beigeladenen“ rechtfertige „den Rückgriff auf das Hilfskriterium der Frauenförderung (noch) nicht.“ Der Leistungsgrundsatz dürfe nicht relativiert werden.[49] Diese Rechtsprechung zwingt zu einer Bewertung der Einzelkriterien. Es liegt nahe, dass ein Beurteilungsgleichstand bei einer Vielzahl von Beurteilungskriterien kaum jemals festzustellen ist. Nach *Lorses* Analyse der Rechtsprechung läuft die Frauenquote damit praktisch ins Leere.[50] Der Herstellung von Geschlechtergerechtigkeit im Wege einer Frauenquote kann also nur durch eine Reduzierung der Beurteilungskriterien auf ein notwendiges Maß erreicht werden.

Auch bei gebotenen Zweifeln an manchen Instrumenten besteht doch noch ein großes ungenutztes Potenzial zur Verbesserung der Geschlechtergerechtigkeit im Beurteilungswesen durch Veränderung von Verfahrensweisen.

2. Sachlichkeit bei den Beurteilungskriterien

Um den beschriebenen inhaltlichen Hindernissen für ein geschlechtergerechtes Beurteilungswesen zu begegnen, ist zunächst in Auswertung des Evaluationsgutachtens zum FüPoG eine zahlenmäßige Reduzierung der Beurteilungskriterien erforderlich. Durch den damit einhergehenden höheren Abstraktionsgrad können auch Unterschiede in verschiedenen Aufgabenbereichen berücksichtigt werden, so dass die Beurteilungsrichtlinien allgemeiner anwendbar sind.

Männlich attribuierte Eigenschaften (wie Durchsetzungsfähigkeit) und sekundäre Leistungen (wie mündlicher und schriftlicher Ausdruck) sollten keinen Platz mehr in Beurteilungsrichtlinien haben. Stattdessen empfiehlt es sich Ansätze zu prüfen, die die paternalistischen Strukturen der bisherigen Beurteilungspraxis aufbrechen, indem sie einen frauenfreundlicheren und gleichzeitig leistungsorientierteren Ansatz fördern. Mit einer reinen Ergebnis-Orientierung allein ist es damit

[48] 73. DJT (2022 in Bonn) Beschlüsse 20 und 19 (https://djt.de/wp-content/uploads/2022/09/Beschluesse.pdf).
[49] OVG NRW, Beschl. V. 21.2.2017, 6 B 1109/16, juris, Rn. 53, https://openjur.de/u/966121.html, Rn. 53 und Rn 76.
[50] *Lorse* (2017), insbes. S. 460-464; 461 m. Verweis auf Rechtsprechung der VG Düsseldorf und Aachen und 462.

allerdings nicht getan. Allein Menge der Entscheidungen und Geschwindigkeit, mit der sie erreicht werden, gewährt weder Qualität noch Verbindlichkeit oder gar eine angemessene Fehlerkultur und Evaluation des Vorgehens. Für eine verantwortliche Entscheidungskultur empfiehlt es sich daher, in jeweils der Sache angemessenem Maß die Schritte zu beachten, die für ein rechtsstaatliches Verwaltungshandeln erforderlich sind, nämlich (1) die Ermittlung des Sachverhalts, wobei von Führungskräften erwartet werden muss, dass sie proaktiv auf ihre Mitarbeitenden zugehen, um von möglichen Problemen zu erfahren, (2) die Würdigung des Sachverhalts, (3) die Entscheidungsfindung und (4) deren Umsetzung. Durch ein derartiges Eingehen auf die bewährten Schritte der Entscheidungsfindung in der Verwaltung wird statt eines Perfektionsanspruchs, der Entscheidungen oft hemmt, eine positive Kultur gefördert, die Beiträge zu Erfolgen (output) leistet. Dabei gehört das frühzeitige Erkennen von Herausforderungen ebenso zur Aufgabenerfüllung, wie die Analyse des Problems, die Entscheidungsfindung und die Umsetzung der Lösung. Wichtiger als die aus der paternalistischen Tradition stammende Aufzählung einzelner Aufgaben, die von einer höheren Hierarchiestufe übertragen werden, ist die selbständige Wahrnehmung eines umfassenden Verantwortungsbereichs. Ein derartiges auf Diskurs statt Dominanz ausgerichtetes Leistungsverständnis eröffnet Frauen, wie oben dargelegt, eher die Möglichkeit sich zu profilieren und ist, wie ebenfalls bereits hervorgehoben, auch ein Beitrag im öffentlichen Interesse der Leistungserbringung. Positiv unterstellt, entspricht es auch dem, was die neue Bundesregierung mit ihren Zielen von „Kulturwandel und moderne[r] Führung“, öffentlichen Dienst attraktiver machen und Modernisierung des öffentlichen Dienstrechts meint.[51]

D. Zusammenfassung

Der djb schlägt daher vor, gerade im Interesse von Frauen, andere Aspekte des Koalitionsvertrags für die Bestenauslese zu betrachten, als sie im Koalitionsvertrag zur Frauenförderung benannt sind. Konkret sieht der djb drei Schwerpunkte.

Zum Ersten bedarf es dringend einheitlicher und transparenter Beurteilungsstandards in der Bundesverwaltung als Grundlage für ein geschlechtergerechtes Beurteilungswesen. Der djb empfiehlt die Festlegung einer zentralen Zuständigkeit in der Geschäftsordnung der Bundesregierung.

Zum Zweiten bedarf es einer geschlechtergerechten Ausgestaltung des Beurteilungsverfahrens, um die Anwendung sachfremder Erwägungen zu vermeiden. Hierzu schlägt der djb den Ausschluss von Personen aus der Beurteilendenfunktion vor, die selbst ihr Amt nicht auf Grund eines Auswahlverfahrens erhalten haben. Anstelle eines mehrstufig hierarchischen Beurteilungsverfahrens, das auf einzelne Personen ausgerichtet ist, sind paritätisch besetzte Beurteilungskommissionen zu schaffen.

Zum Dritten ist der Katalog von Beurteilungskriterien deutlich zu reduzieren, insbesondere um männlich attribuierte Eigenschaften (wie Durchsetzungsfähigkeit, Belastbarkeit) und sekundäre Leistungen (wie mündlicher und schriftlicher Ausdruck). Stattdessen empfiehlt sich eine Kriterienstruktur, die an Verantwortungswahrnehmung statt Perfektion von Einzelergebnissen orientiert ist.

Für das Fortkommen von Frauen im öffentlichen Dienst sind Gleichbehandlung und Leistungsorientierung im Beurteilungswesen fundamentale Voraussetzungen, die derzeit nicht ausreichend gewährleistet sind.

[51] https://www.koalitionsvertrag2025.de/sites/www.koalitionsvertrag2025.de/files/koav_2025.pdf, Zeilen 1834 ff., 1841 ff., 1846 ff.

Impressum

Herausgeber: Deutsche Juristinnenbund e.V. (djb)

Präsidium: Ursula Matthiessen-Kreuder, Rechtsanwältin, Bad Homburg (Präsidentin), Dr. Lucy Chebout, M.A., Rechtsanwältin, RinVerfGH Berlin und Verena Haisch, Rechtsanwältin, Hamburg (Vizepräsidentinnen), Petra Lorenz, Regierungsdirektorin i.R. (Schatzmeisterin), Anke Gimbal, Rechtsassessorin, Berlin (Geschäftsführerin)

Verantwortlich: Dr. Marianne Czisnik (stellvertretende Vorsitzende der djb-Kommission Verfassungsrecht, Öffentliches Recht, Gleichstellung)

Bundesgeschäftsstelle
Kronenstr. 73
10117 Berlin

Telefon: +49 30 4432700

https://www.djb.de/
geschaeftsstelle@djb.de

AG Dortmund, Vereinsreg.-Nr.: 1444

Berlin, 2025