

Diskriminierungsrisiken beim Zugang zum Polizeidienst

Die Auswahlverfahren der Polizeien zwischen Bestenauslese und Benachteiligung

Sabrina Ellebrecht

1. Einleitung

Welche Gründe gibt es für die Unterrepräsentation von Migrant:innen in den deutschen Polizeien? Und welche Rolle spielen dabei die Auswahlverfahren für den Polizeidienst? Kommt es im Rahmen der Personalauswahl zu benachteiligenden Selektionen gegenüber Bewerber:innen mit Migrationshintergrund? Diese Fragen waren leitend für die Forschungsarbeiten zum Umgang der Polizeien mit Diversität in der Personalauswahl, deren Ergebnisse hier zusammenfassend dargestellt werden.¹ Sie knüpfen an die Forderung der Bundesregierung aus dem Jahr 2014 an, »mögliche mittelbare Diskriminierungen« gegenüber dem »Bewerberkreis mit Migrationshintergrund« auszuschließen und »in den Einstellungsverfahren sprachliche und interkulturelle Kompetenzen angemessen« (Deutscher Bundestag 2014: 14) zu berücksichtigen.

Mit dem Verweis auf »mögliche mittelbare Diskriminierungen« hat die Bundesregierung eine weitreichende Formulierung gewählt. Denn wenn Regelungen oder Kriterien, die für alle in gleicher Weise gelten – etwa die Zulassungsbedingungen und Prüfkriterien für den Polizeidienst –, in ihrem Effekt einzelne Gruppen stärker benachteiligen als andere, besteht bereits die Vermutung einer Diskriminierung.²

1 Ein herzlicher Dank gilt Christina Bott, die die Datenerhebung und -auswertung für dieses Teilprojekt als wissenschaftliche Hilfskraft von 2020 bis 2021 unterstützt hat. Sie hat in einem Bundesland einige der teilnehmenden Beobachtungen der Auswahlgespräche übernommen sowie die statistischen Berechnungen umgesetzt.

2 Die Antidiskriminierungsstelle des Bundes (ADS) definiert mittelbare Diskriminierung als Benachteiligung einer Person, die »nicht offensichtlich wegen eines in § 1 AGG genannten Merkmals (erfolgt), sondern [...] aus scheinbar neutralen Kriterien (resultiert). Diese gelten zunächst für alle gleichermaßen, in ihrem Effekt aber wirken sie sich auf bestimmte Gruppen stärker benachteiligend aus als auf andere« (<https://www.antidiskriminierungsst.de>

In diesem Fall ist sicherzustellen, dass berufsbezogene Anforderungen die Selektion begründen und nicht etwa Herkunft oder andere durch den Gleichbehandlungsgrundsatz geschützte Kategorien systematisch relevant werden. Auswahlverfahren sollen selektieren, dürfen aber nicht diskriminieren.

Der Beitrag widmet sich den Personalauswahlverfahren deutscher Polizeibehörden und analysiert die Zugangschancen für migrantische Bewerber:innen sowie die Zugangshürden, die spezifisch zu ihren Lasten gehen. Ziel des Beitrags ist es, jene Elemente der Personalauswahl zu identifizieren, die verstärkt migrantische Bewerber:innen herausselektieren. Dazu wurden einerseits Erfolgsquoten und Testergebnisse von Bewerber:innen mit und ohne Migrationshintergrund verglichen und andererseits Bewertungsvorgänge analysiert. Die Analysen sind exemplarisch und ziehen nicht auf landes- oder behördenspezifische Aussagen. Im Fokus stehen Diskriminierungsrisiken, die mit charakteristischen Testelementen und Anforderungen verbunden sind und deshalb für die Auswahlverfahren aller Polizeien zu überprüfen wären.

2. Vielfalt und Bestenauslese: ein Widerspruch?

Seit etwa 15 Jahren wird in der politischen Öffentlichkeit der prozentuale Anteil neu eingestellter Polizist:innen mit Migrationshintergrund als Indikator für Vielfalt in den Polizeien herangezogen. Dazu beigetragen haben insbesondere die Rechercheberichte des Netzwerks Mediendienst Integration, die seit 2014 regelmäßig veröffentlicht werden. In drei Bundesländern entsprechen die Anteile neueingestellter Polizist:innen mit Migrationshintergrund seit einigen Jahren ungefähr dem Bevölkerungsanteil, das sind in Berlin und Baden-Württemberg gut 30 Prozent und in Sachsen-Anhalt 6,5 Prozent. In den übrigen Landespolizeien sind Menschen mit Migrationshintergrund bei den Neueinstellungen weiterhin unterrepräsentiert (vgl. Mediendienst Integration 2022).

In diesem Zusammenhang ist es wichtig zu unterscheiden: Die Frage nach den Zugangshürden auf Seiten der Polizeiorganisation für migrantische Bewerber:innen spricht andere Ausschlussmechanismen an als die institutionellen Barrieren, die in den 1990er Jahren diskutiert wurden, als von ›Interkultureller Öffnung‹ und einer ›multiethnischen Polizei‹ die Rede war. Seinerzeit ging es um institutionelle, konkret um beamtenrechtliche Hürden. Zur Diskussion stand eine Lockerung des Staatsangehörigkeitserfordernisses. Es ging um ›Ausländer in der Polizei‹. Seit Anfang der 2000er Jahre geht es um die weitaus komplexere Frage, inwieweit Zuge-

elle.de/DE/ueber-diskriminierung/was-ist-diskriminierung/diskriminierungsformen/diskriminierungsformen-node.html [Zugriff: 18.09.2023]). Die Polizeien unterliegen in der Rolle der Arbeitgeberin dem Allgemeinen Gleichbehandlungsgesetz (vgl. AGG § 24).

wanderte und ihre Nachkommen (chancen-)gleichen Zugang zum Polizeidienst als öffentlichem Amt haben. Diese sind mehrheitlich deutsche Staatsbürger, teils mit doppelter Staatsangehörigkeit, oft in Deutschland geboren und haben keine eigene Migrationserfahrung. Sie sind reguläre Adressaten von Art. 33 Abs. 2 GG der maßgeblichen Norm für den Zugang zum Polizeidienst. Die Unterrepräsentation ist also erklärungsbedürftig. Die Zielgruppe migrantischer Bewerber:innen kann bspw. nicht informiert, nicht interessiert oder nicht qualifiziert sein, oder die Polizei generell ablehnen, so dass der Beruf für sie nicht in Frage kommt. Diese Faktoren der Selbstselektion kommen allerdings während des Auswahlverfahrens kaum zum Tragen, so dass hier die behördenseitigen Selektionsmechanismen im Fokus stehen.

Für eine Analyse der Zugangshürden gegenüber migrantischen Bewerber:innen stelle ich zunächst einige Überlegungen zur Norm des gleichen Zugangs und zur Bestenauslese voran (2.1). Danach wird das Auswahlverfahren beschrieben. Es werden einzelne Schritte und Testelemente identifiziert, die als Filter fungieren und grundsätzlich zum Ausschluss aus dem Verfahren führen können (2.2). Vor der Darstellung der empirischen Ergebnisse, die Diskriminierungsrisiken des PC-Tests und des kommunikativen Testteils fokussieren, stelle ich das methodische Vorgehen und das erhobene Material vor (2.3).

2.1 Gleicher Zugang

Der Zugang zum Polizeidienst soll ein gleicher sein, so gibt es Art. 33 Abs. 2 GG vor. Darin ist formuliert, dass jede:r Deutsche »nach seiner Eignung, Befähigung und fachlichen Leistung gleichen Zugang zu jedem öffentlichen Amte« hat. Das Auswahlverfahren für den Polizeidienst soll diesen »gleichen Zugang« umsetzen, und zwar für einen über das Erfordernis der Staatsangehörigkeit definierten Personenkreis³ und indem all diese an den gleichen Anforderungen gemessen und nach dem Leistungsprinzip bewertet werden. In den Auswahlverfahren werden die Bewerber:innen anhand identischer Kriterien in einem standardisierten Verfahren durch Punkte bewertet. Anhand der Punktewerte werden sie dann in eine Rangliste gebracht, und die Kandidat:innen auf den oberen Plätzen bekommen die freien Stellen. Gleiche Anforderungen und gleiche Prüfungsbedingungen befördern prinzipiell gleichen Zugang, so die Grundidee.

Allerdings gibt es Zweifel daran, ob Gleichbehandlung als identische Behandlung im Verfahren der Norm des gleichen Zugangs genügt. Nach Daniela Hunold

3 Der Zugang zum Polizeidienst war lange exklusiv an die deutsche Staatsangehörigkeit geknüpft. Der sogenannte Deutschenvorbehalt ist mittlerweile durch die Arbeitsmarktfreizügigkeit von EU und EWG aufgehoben. Für Staatsangehörige von Nicht-EU-Staaten, oft Drittstaatenangehörige genannt, gelten jeweils spezifische Regelungen (vgl. Tabbara 2013; Ellebrecht 2022b).

(2008:46) ist der »dogmatische Bezug auf die Gleichheitsidee ein Kriterium, das Ungleichheit produziert, indem Unterschiede negiert und Zugangschancen faktisch ungleich werden«. Auch Rafael Behr und Annelie Molapisi (2022: 71) weisen darauf hin, dass es die oberste Priorität der Polizeiorganisation sei, »alle formal gleich zu behandeln«. Die »nicht-kompatible Vielfalt«, das sind jene, die sich nicht zur Eingliederung eignen, würden »an der Pforte (d. h. im Auswahlverfahren)« abgewiesen (ebd.: 72). Bestenauslese bezieht sich danach nicht primär auf die Eignung für den Beruf, sondern vielmehr auf die Passung zur Organisationskultur. Das führt im Ergebnis dazu, dass die Ähnlichen für die Besten gehalten werden.

Ihrer Grundidee nach trägt die Bestenauslese jedoch zur gesellschaftlichen Fairness bei, indem Leistungskriterien und nicht soziale Beziehungen die Stellenbesetzung regulieren (vgl. Görlich/Schuler 2014: 1181; Leenen et al. 2021: 95). Ein Rekrutieren der Ähnlichen, wie es bei Behr und Molapisi anklingt, ist im Ergebnis nicht unbedingt eine Auslese der Besten für den Beruf, d. h. für die vielen unterschiedlichen Aufgaben und Publikumskontakte der Polizeiarbeit. Auch rechtshistorisch, so stellt Laura Wisser (in diesem Band: 127) heraus, war das Gebot der Bestenauslese »vor allem eine Absage an Nepotismus und Ämterkäuflichkeit«. Durch Leistungsorientierung sollte nicht die übliche Klientel, sondern es sollten die Besten aus allen Milieus und Gruppen rekrutiert werden. Ein Auswahlverfahren, in dem einzelne gesellschaftliche Gruppen systematisch schlechter abschneiden, trägt dem Grundgedanken der Bestenauslese also nicht Rechnung. Vielmehr kann umgekehrt davon ausgegangen werden, dass diskriminierende Strukturen im Auswahlverfahren die Bestenauslese kompromittieren. Deshalb kann für die Norm des gleichen Zugangs zusätzlich ein Blick auf Chancengleichheit relevant sein.

2.2 Die Auswahlverfahren der Polizei

Zuständig für die Personalgewinnung der Polizeien sind die Werbe- und Einstellungsstellen der jeweiligen Landespolizeien bzw. des Bundeskriminalamtes und der Bundespolizei.⁴ Sie führen die Auswahlverfahren durch, und planen und gestalten außerdem die Werbe- und Informationsmaßnahmen. Bis zur Einstellung ins Beamtenverhältnis auf Widerruf, dem Beginn von Ausbildung oder Studium, führen sie die Akten der Bewerber:innen. Durch die Werbemaßnahmen tragen diese Funktionseinheiten dazu bei, dass der Bewerberpool möglichst groß und durchmischt ist. Außerdem haben sie die Aufgabe, aus dem Pool wiederum zu selektieren: erstens durch Überprüfen der *Formalkriterien* in den eingegangenen Bewerbungsunterlagen; zweitens durch *Bestenauslese* im Rahmen des Auswahlverfahrens; und drittens durch die grundsätzliche Überprüfung der *Eignung*, die innerbehördlich in der Regel

4 Das Verfahren habe ich in ähnlicher Form auch an anderer Stelle beschrieben (vgl. Ellebrecht 2022b). Dort gebe ich weiterführende Hinweise zum Begriff der Eignung.

als charakterliche Eignung spezifiziert wird und durch unterschiedliche Elemente im Verfahren überprüft wird. Neben den Persönlichkeitstests wird die charakterliche Eignung auch über die Zuverlässigkeitstests und ggf. eine Sicherheitsüberprüfung festgestellt. Es dürfen keine Vorstrafen oder laufende Verfahren mit den Bewerbern verbunden sein, auch können Falschaussagen oder Verschweigen während des Bewerbungsverfahrens mit dem Hinweis auf mangelnde charakterliche Eignung zum Ausschluss führen. Schließlich wird im Laufe des Auswahlverfahrens die sogenannte Polizeidiensttauglichkeit durch den Polizeiärztlichen Dienst überprüft. Zur Anwendung kommt dabei die Polizeidienstvorschrift (PDV) 300. Die Befunde des Polizeiärztlichen Dienstes werden in einer separaten Akte geführt.⁵

Jede einzelne Stufe des Auswahlverfahrens, die zum Ausschluss aus dem Verfahren führen und nicht durch andere Tests ausgeglichen werden kann, kann als Hürde identifiziert werden. Dazu gehören neben den Vorauswahl- bzw. Formalkriterien, kognitive Leistungstests, ein Sporttest, ein persönliches Gespräch bzw. die kommunikativen Aufgaben im Assessment-Center, die Überprüfung durch den Polizeiärztlichen Dienst und die Zuverlässigkeitstests. Wer auch nur über eine Stufe ›stolpert‹, wird vom weiteren Verfahren ausgeschlossen, unabhängig davon, ob er oder sie in den anderen Tests eine hohe Punktwertung erreicht hat. Das Auswahlverfahren findet entweder an einem oder an zwei Tagen statt und besteht aus drei Teilen: schriftlicher Leistungstest, Sportprüfung und kommunikativer Test. Die schriftliche Prüfung und der Sporttest werden ausnahmslos vor dem mündlichen Teil durchgeführt. Die Tauglichkeitsprüfung findet in einigen Bundesländern vor und in anderen nach dem kommunikativen Teil statt. Bevor das Auswahlverfahren startet, müssen die grundsätzlichen Voraussetzungen für den Polizeiberuf und das Berufsbeamtentum geklärt werden. Zu den Bewerbungsvoraussetzungen gehören neben dem Staatsangehörigkeitserfordernis definierte Formalkriterien wie Alter, Körpergröße⁶ und BMI sowie der Schulabschluss. Diese Kriterien stellen institutionelle Hürden dar und sind in verschiedenen Geset-

-
- 5 Die Entscheidungen des Polizeiärztlichen Diensts sind in dieser Untersuchung eine *black box* geblieben. Für das Fortkommen im Auswahlverfahren zählt die Information, ob ein:e Bewerber:in polizeidiensttauglich ›geschrieben wurde‹ oder nicht. Die Tests des Polizeiärztlichen Dienstes fallen damit als Ja-Nein Entscheidung im Auswahlverfahren ins Gewicht. Untauglich heißt Ausschluss aus dem Verfahren. Die Inhalte der Untersuchung sind nicht Gegenstand der Bewerberakte. In mehreren Situationen bei Zweifel an der charakterlichen Eignung oder sozialen Passung habe ich die Formulierung, dass der Polizeiärztliche Dienst ›da nochmal draufschauen sollte‹ notiert, so dass bei mir der Eindruck entstanden ist, dass diese Ja-Nein Entscheidung den Spielraum für einen Ausschluss bei Zweifeln an einer Person öffnet.
- 6 Mittlerweile wurde in einigen Bundesländern sowie bei der Bundespolizei eine Mindestkörpergröße als Zulassungskriterium fallen gelassen.

zen und Verordnungen geregelt, die je nach Bundesland variieren können.⁷ Das mehrstufige Auswahlverfahren überprüft dann Eignungs- und Leistungskriterien. Teilweise findet die Bepunktung dieses hochstandardisierten Massenverfahrens automatisiert statt. Bei den PC-Tests etwa ergeben sich die Testergebnisse direkt als Punktwert. Eine Korrektur der Aufgaben durch Personaler der Polizei wird kaum noch vorgenommen.

Die Auswahlverfahren stellen einen zentralen, regulativen Punkt für den Zugang zum Polizeidienst dar. Im Vergleich zur weitgehend diffus adressierten Anwerbung findet an dieser Stelle aktive Selektion durch die Organisation statt. Bei den Auswahlverfahren steht die Bewertung und Auswahlentscheidung der Organisation im Vordergrund.

2.3 Material und Methoden

Um – wie von der Bundesregierung angeregt – »mögliche mittelbare Diskriminierungen« gegenüber dem »Bewerberkreis mit Migrationshintergrund« auszuschließen, muss es zunächst möglich sein, Aussagen über diesen Bewerberkreis zu treffen. Die Voraussetzung dafür ist, dass die Kategorie »mit Migrationshintergrund« im Rahmen der Bewerbung erfasst wird. Dass dies nicht bundesweit und auch nicht einheitlich geschieht, verdeutlichen verschiedene Forschungsarbeiten und Rechercheberichte, die sich einer Quantifizierung von Vielfalt in der Polizei angenommen haben.⁸ Nur Polizeien, die die Variable »Migrationshintergrund« zu Beginn des Ver-

-
- 7 Jedes Bundesland hat spezifische Laufbahnverordnungen, die die Voraussetzungen für die Einstellung in den Polizeidienst festlegen. Diese Verordnungen regeln unter anderem die erforderlichen Schulabschlüsse, die gesundheitlichen Anforderungen und das Auswahlverfahren. Das Beamtenstatusgesetz (BeamtStG) hingegen regelt, welches die allgemeinen Voraussetzungen für die Einstellung ins Berufsbeamtentum des öffentlichen Dienstes sind, einschließlich der Polizei. Teilweise enthalten auch die Polizeigesetze der Länder Bestimmungen zu den Anforderungen an Polizeibeamte und -beamtinnen.
- 8 Während Maurer 1995 und Maguer 2002 noch die Anzahl ausländischer Beamter dokumentieren, finden sich bei Hunold (2008: 29–34) erste Informationen zu Anteilen deutscher Polizisten mit Migrationshintergrund. Die Angaben dazu stammen aus einer entsprechenden Bund-Länder-Abfrage im Rahmen des Projekts »Migranten in Organisationen von Recht und Sicherheit (MORS)«. 2013 gelingt es Regge im Rahmen einer Diplomarbeit den Migrantenanteil der Neueinstellungen (2007 bis 2012) für den Polizeivollzugsdienst eines Großteils der Länder sowie des Bundes darzustellen. Der Mediendienst Integration hat seit 2014 insgesamt sechs Rechercheberichte veröffentlicht, in denen er den Anteil neueingestellter Polizist:innen mit Migrationshintergrund zum Anteil der Bevölkerung im Bundesland vergleicht. Eine Zusammenfassung der ersten fünf Berichte habe ich an anderer Stelle formuliert (vgl. Ellebrecht 2002b: 528–530). Bemerkenswert scheint mir, dass trotz öffentlicher Nachfragen, die Einstellungsstellen noch nicht dazu übergegangen sind, die Zahlen selbst zu veröffentlichen oder einheitliche Erhebungen umzusetzen.

fahrens erheben, können Erfolgsquoten als Verhältnis von Neueinstellungen zu Bewerbungen ermitteln und vergleichen.⁹ Die Datenerfassung der Bewerberverwaltung bedingt also die Möglichkeit für Untersuchungen von benachteiligenden Effekten in den Auswahlverfahren.¹⁰

Die Erfassung einer gruppenspezifischen Benachteiligung kommt um die Verwendung der Differenzkategorie nicht herum. »Migrationshintergrund« wird hier lediglich als statistische Hilfsvariable verwendet, der Bewerberkreis mit Migrationshintergrund wird statistisch mit jenem ohne Migrationshintergrund verglichen.

Deshalb sollte im ersten Schritt geklärt werden, ob und in welcher Weise die Bewerberverwaltungen personelle Vielfalt und die Kategorie »Migrationshintergrund« erheben. Dazu wurde im Sommer 2019 im Rahmen einer allgemeinen *Datenanfrage* zur Zusammensetzung von Bewerbungen und Neueinstellungen seit 1990 auch die Erfassungsweise der Behörden erfragt. Alternativ zur Übermittlung von aggregierten Daten wurden Jahresberichte der Einstellungsstellen seit den 1990er Jahren erbeten, um Datenerfassung und Trends gegebenenfalls zu rekonstruieren. Die Daten, die in aggregierter Form übermittelt wurden, waren hinsichtlich Vergleichbarkeit und Datendichte kaum ergiebig. Allerdings konnten anhand der Jahresberichte die bürokratischen *work-arounds* bei der Erfassung von Vielfalt nachgezeichnet werden (vgl. Ellebrecht 2022b). So nutzte die Bewerberverwaltung der Polizei Hamburg die Bezeichnung »Migrationshintergrund« erst seit 2012. Zuvor gab es die Unterscheidung »Bewerber ausländischer Herkunft«, jeweils unterteilt in »ohne deutsche Staatsangehörigkeit« und »mit deutscher Staatsangehörigkeit« (ebd.: 525). Brandenburg führt die Kategorie »Muttersprachler«, wobei dies sowohl »ausländische Staatsbürger, als auch Deutsche mit einer zweiten Muttersprache, also einem Migrationshintergrund, sein« können (ebd.). Zwar wird oft betont, dass beamtenrechtlich nur die Staatsangehörigkeit von Relevanz sei. Der Blick der Verwaltung markiert aber neben der klaren Unterscheidung von Staatsangehörigkeit Abstufungen von Zugehörigkeit, indem sprachliche und kulturelle Unterscheidungsmerkmale erfasst werden, die im Polizeialtag nützlich sein könnten.

Ebenfalls Teil der ersten Forschungsanfrage war die Durchführung von *teilnehmenden Beobachtungen* während des Auswahlverfahrens, da durch sie »Aspekte des Handelns und Denkens beobachtbar werden, die in Gesprächen und Dokumenten – gleich welcher Art – über diese Interaktionen bzw. Situationen nicht in dieser

9 Das waren zum Zeitpunkt meiner Datenerhebung die Einstellungsstellen in Berlin, Bremen, Hamburg, Niedersachsen und Nordrhein-Westfalen.

10 Der Verweis auf fehlende Daten seitens der Behörden, wie etwa von einem Experten im Interview mit Kreyenborg formuliert (in diesem Band: 67–85), ist also in den Verantwortungsbereich der Behörden selbst zurückzugeben.

Weise zugänglich wären« (Lüders 2003: 151). Teilnehmende Beobachtungen waren also das Mittel der Wahl und nicht durch Experteninterviews zu ersetzen. So konnten neben den standardisierten Vorgaben auch emotionale, körperliche, habituelle oder stimmliche Eindrücke gewonnen werden. Durch die teilnehmende Beobachtung in Auswahlsituationen war es möglich, die Diskrepanz zwischen der Routine des Massenverfahrens aufseiten der Bewertenden und der persönlichen Prüfsituation der Bewerber:innen mitzuerleben sowie einen Eindruck wechselseitiger Sympathien und Irritationen zu bekommen. Für den kommunikativen Tests war ich darauf angewiesen, dass eine Zuordnung zur Gruppe »mit Migrationshintergrund« in der Interaktion vorgenommen wurde, und zwar entweder durch eine solche Formulierung durch den oder die Bewerber:in selbst oder durch eine entsprechende Zuschreibung durch die Bewertenden. Auf ein eigenes ›Lesen‹ eines möglichen Migrationshintergrunds habe ich verzichtet.

Zwischen Februar 2020 und Juni 2021 war ich in vier Landespolizeien bei den Auswahlverfahren vor Ort bei den schriftlichen Tests, Sporttests und den mündlichen Auswahlgesprächen dabei. Bei den persönlichen Interviews hatte ich keinen Einfluss auf das Sample, sondern beobachtete und protokollierte die Auswahlgespräche wie sie vorgesehen waren ›der Reihe nach‹. So konnte ich 32 persönliche Auswahlgespräche erfassen, in denen 22 männliche und 10 weibliche Bewerber:innen geprüft wurden. Von denen schieden insgesamt 8 nach dem Interview aus. Bei fünf Personen wurde während des Auswahlgesprächs ein Migrationshintergrund thematisiert.¹¹

Im zweiten Schritt wurden bei den Einstellungsstellen der Polizei Berlin, Hamburg, Hessen, Rheinland-Pfalz und der Bundespolizei *anonymisierte Individualdaten* zu den letzten Bewerberjahrgängen angefragt, aus denen die Punktwerte der Bewerber:innen pro Testelement ersichtlich wären. Die Beantwortung dieser Anfragen war langwierig. Wiederholt wurde ich vertröstet oder um weitere Unterlagen und Stellungnahmen gebeten. Wechselnde Zuständigkeiten in den Behörden machten zum Teil erneute Anfragen notwendig und führten in einem Fall auch zur Zurücknahme der Kooperationsbereitschaft.¹² Die Forschungskooperationen mit Berlin und Rheinland-Pfalz wurden schließlich dadurch realisiert, dass die statistische Auswertung

¹¹ Eine Kandidatin antwortete auf die Frage, ob sie wisse, wofür die AfD als Partei stehe, dass sie »als Ausländerin wisse, dass die Partei gegen sie sei«. Diese Selbstverortung ist erwähnenswert, da die Kandidatin laut Lebenslauf die deutsche Staatsangehörigkeit hat, in Deutschland geboren ist und zweisprachig aufwuchs.

¹² Es wurde die Freigabe von Datenmaterial, das über einen Zeitraum von sechs Monaten zusammengetragen worden war, zurückgezogen. Die Daten waren bereits auf einem verschlüsselten Stick bei der Behörde hinterlegt. Als es darum ging, den Termin zu vereinbaren, zu dem wir diesen Stick persönlich abholen könnten, wurde die Freigabe der Daten zurückgezogen.

der relevanten Daten vor Ort in den Räumlichkeiten der Behörden durchgeführt wurde. In Rheinland-Pfalz resultierte aus dem Termin eine Kooperation für eine gemeinsame Studie.¹³ Die Auswertungen für Berlin wurden in dem Zeitraum eines Nachmittags durchgeführt, mit allen damit verbundenen Nachteilen, bspw. konnten Kontroll- oder ergänzende Auswertungen nicht stattfinden. Diese statistischen Auswertungen orientierten sich an dem Vorbild der Studie von Stumpf et al. (2016) und zielen darauf ab, (1) das Chancenverhältnis von Bewerber:innen mit und ohne Migrationshintergrund zu bestimmen (*adverse impact*) und (2) einen Mittelwertvergleich der Testergebnisse dieser Gruppen durchzuführen (Varianzanalyse, Cohen's d).¹⁴ Folgende Stichproben wurden analysiert:

- Bewerberjahrgang 2020 für den Bachelorstudiengang »Polizeidienst« an der Hochschule der Polizei Rheinland-Pfalz mit $N_1 = 3005$ Bewerber:innen gesamt; vorausgesetzt ist das Abitur oder die Fachhochschulreife, denn in Rheinland-Pfalz gibt es nur den gehobenen Dienst (gD).
- Bewerberjahrgang 2020 für den Seiteneinstieg in die Polizei Rheinland-Pfalz mit $N_2 = 289$; vorausgesetzt ist ein Realschulabschluss. (Die Ausbildungsjahre bereiten auf das Polizeistudium vor, für das man sich dann nicht mehr bewerben muss.) Es ist ein Seiteneinstieg in den gehobenen Dienst.
- Stichprobe aus der Bewerberkohorte Herbst 2020 der Berliner Polizei mit $N_3 = 6026$. In Berlin kann man sich mit Abitur für den gehobenen Dienst (gD) bei der Kriminalpolizei (Kripo) oder der Schutzpolizei (Schupo) bewerben. Es ist auch möglich, sich mit Abitur für den mittleren Dienst der Berliner Schutzpolizei zu bewerben, ein Realschulabschluss ist Grundvoraussetzung. Bewerber:innen, die sich in der Rubrik »Lebensältere« im mittleren Dienst beworben haben, wurden für die Stichprobe nicht berücksichtigt.
- Datensatz zu den Bewerbungen bei der Berliner Polizei für die Jahre 2006–2020; darin sind aggregierte Informationen, bspw. bestanden: ja/nein, und keine Punktewerte festgehalten.

Ergänzend führte ich elf *teil-strukturierte Interviews* mit insgesamt 14 Verantwortlichen der Personalauswahl, Bewertenden, Eignungsdiagnostiker:innen und Arbeits-

¹³ Die Adverse Impact Studie Rheinland-Pfalz habe ich zusammen mit Christina Bott und Markus Thielgen durchgeführt. Die Publikation steht noch aus.

¹⁴ Die Berechnungen wurden mit SPSS durchgeführt, mit Ausnahme der Berechnung von Cohen's d; dieser Wert wird von SPSS nicht ausgewiesen und muss separat berechnet werden. Dazu wird der Differenzwert der Mittelwerte der beiden Vergleichsgruppen durch die mit den Stichprobenumfängen gewichtete Standardabweichung (gepoolte Standardabweichung) geteilt (vgl. Stumpf et al. 2016: 4).

psycholog:innen.¹⁵ Einige Interviews wurden *face-to-face* in den Räumlichkeiten der Einstellungsstellen durchgeführt, andere telefonisch. Von drei Interviews wurden Gesprächsprotokolle angefertigt, acht konnten mitgeschnitten und danach transkribiert werden. In den Gesprächen ging es um die jeweils landesspezifischen Auswahlverfahren, generelle Rekrutierungsschwierigkeiten der Polizeien, die Einschätzung der aktuellen Bewerbergeneration sowie die Bedeutung von Diversität und Bewerber:innen mit Migrationshintergrund für die Polizei und ihre Personalentwicklung. Die Fragen wurden offen und erzählgenerierend gestellt. Dabei verstand ich die Befragten »nicht als ›Experten‹, die mir ›objektive‹ Informationen zur Verfügung stellten, sondern als Akteure im Feld, die Deutungsmacht beanspruchen und an der Herstellung und Verbreitung von Deutungen aktiv mitwirken« (Lang 2019: 105-106).

Das hier beschriebene Material wird in den nächsten beiden Abschnitten nicht vollumfänglich analysiert. Ich beschränke mich stattdessen auf zentrale Ergebnisse, in denen ungleiche Bewertungen oder überproportionale Selektionen zulasten von Bewerber:innen mit Migrationshintergrund deutlich wurden.

3. Zugangshürden: Statistische Auffälligkeiten

Statistische Auffälligkeiten liefern Anhaltspunkte für Diskriminierungsrisiken, die einem Verfahren trotz hoher Standardisierung innewohnen. Denn letztlich kumulieren die vielen situativen, organisationalen Selektionsmechanismen und -entscheidungen statistisch im Ergebnis der Personalauswahl. Die kumulativen Effekte werfen also den Scheinwerfer auf jene Stellen im Auswahlverfahren, die ungleiche Verteilungen produzieren. Sie können »Anlass für die Vermutung der direkten oder indirekten Ungleichbehandlung« sein (vgl. Gomolla/Radtke 2009: 126). Ähnlich ist die Bedeutung von Statistik für die Rechtsfigur der mittelbaren Diskriminierung. Statistik hat dabei die Funktion, Effekte zu verdeutlichen, mit denen zunächst eine Diskriminierungsvermutung formuliert werden kann.¹⁶

Das Entscheidende ist, dass diese Auffälligkeiten nicht durch eine Diskriminierungsabsicht oder eine diskriminierende Motivation definiert sind oder erkennbar

15 Interviews mit Bewerber:innen während der teilnehmenden Beobachtung waren nicht möglich, da dies zu stark in den Ablauf eingegriffen hätte. Zudem wiesen die Behörden darauf hin, dass die Bewerber:innen es als Ungleichbehandlung hätten deuten können.

16 Zur Rechtsfigur der mittelbaren Diskriminierung und zu den Schwierigkeiten, den Tatbestand der mittelbaren Diskriminierung festzustellen vgl. die Expertise von Ute Sacksofsky (2010). Zur sozialwissenschaftlichen Definition und Abgrenzung von institutioneller, struktureller, direkter und indirekter Diskriminierung, sowie zur empirischen Diskriminierungs-forschung vgl. den Aufsatz von Mechtild Gomolla (2017).

wären. Denn die (systematische) Benachteiligung indirekter oder mittelbarer Diskriminierung ergibt sich zum Teil durch Konventionen, bestehende Regelungen und Kriterien, die als selbstverständlich und gegeben gelten und teilweise auch als unproblematisch wahrgenommen werden. Als Indikator gelten »erhebliche« Unterschiede. Wie oben beschrieben sind im Folgenden u. a. zwei Größen der analytischen Statistik von Interesse: der *adverse* oder auch *disparate impact* und die Effektstärke Cohen's d. Beide Maße bewerten Unterschiede und bieten Anhaltspunkte, um einen Unterschied als »erheblich« zu klassifizieren. *Adverse impact* bewertet unterschiedliche Erfolgsquoten zweier Gruppen. Allgemein wird von *adverse impact* gesprochen,

»wenn Personalentscheidungen im Rahmen von Einstellungen oder Beförderungen erheblich unterschiedliche Auswahlquoten zum Ergebnis haben, die zulasten der Mitglieder einer bestimmten Rasse, einer ethnischen Gruppe oder eines Geschlechts gehen.« (Stumpf et al. 2016: 2)

Charakteristisch dafür ist, dass es zu nicht intendierte Schlechterstellung einer Gruppe kommt, obwohl das Verfahren standardisiert ist und auf alle gleich angewendet wird. *Adverse impact* wird als relatives Risiko errechnet und ist ein Vergleich der prozentualen Auswahlraten zweier Gruppen.¹⁷ Damit lässt sich das Chancenverhältnis bestimmen und normativ bewerten. Ein *adverse impact* liegt vor, wenn das Verhältnis der Auswahlraten kleiner als 80 Prozent (oder 0,8) ist.¹⁸ Dies wird auch als 4/5-Regel (*four-fifth-rule*) bezeichnet. Mit dieser Faustregel lässt sich potentielle Diskriminierung markieren. Wichtig ist allerdings der Hinweis von Stumpf et al.:

17 In diesem Beispiel errechnet sich der *adverse impact ratio* durch eine Division der Auswahlquoten der beiden Bewerbergruppen. Dabei wird die Erfolgsquote der beiden Gruppen jeweils anteilig zur Anzahl der jeweiligen Bewerbungen gebildet.

18 In der US-amerikanischen Rechtsprechung gilt ein Unterschreiten der 0,8 als erheblicher (substantial) Unterschied und führt zur Umkehr der Beweislast vor Gericht. In Deutschland wird unterstrichen, dass Statistiken nicht verlangt, aber »wertvolle Hilfe für den Nachweis mittelbarer Benachteiligung« sein können (Sacksofsky 2010: 22). Sacksofsky hält die Festlegung einer statistisch diskriminierenden Selektionsquote für zu wenig spezifisch und schreibt: »Wenn Statistiken verwendet werden, muss die unterschiedliche Betroffenheit erheblich genug sein, um eine Diskriminierungsvermutung auszulösen. Freilich ist nicht geklärt, was Erheblichkeit in diesem Sinne genau bedeutet. Man wird wohl kaum mechanistisch eine eindeutige Prozentzahl festlegen können.« (Sacksofsky 2010: 22) Neben der rechtlichen Relevanz ist auch das Qualitätskriterium Testfairness wichtig: Dabei gilt ein Test als fair, d. h. als nicht-diskriminierend, »wenn die resultierenden Testwerte zu keiner systematischen Benachteiligung bestimmter Personen aufgrund ihrer Zugehörigkeit zu ethnischen, soziokulturellen oder geschlechtsspezifischen Gruppen führen« (Moosbrugger/Kelava 2020: 25).

»Tritt Adverse Impact auf, so ist über die Ursachen des Effekts per se noch nichts bekannt: Er kann durch tatsächliche Unterschiede bei Vergleichsgruppen hinsichtlich job-relevanter Merkmale [...] bewirkt werden; er kann aber auch auf verzerrende Vorgehensweisen und Standards im Prozess der Personalauswahl zurückgehen.« (Ebd.)

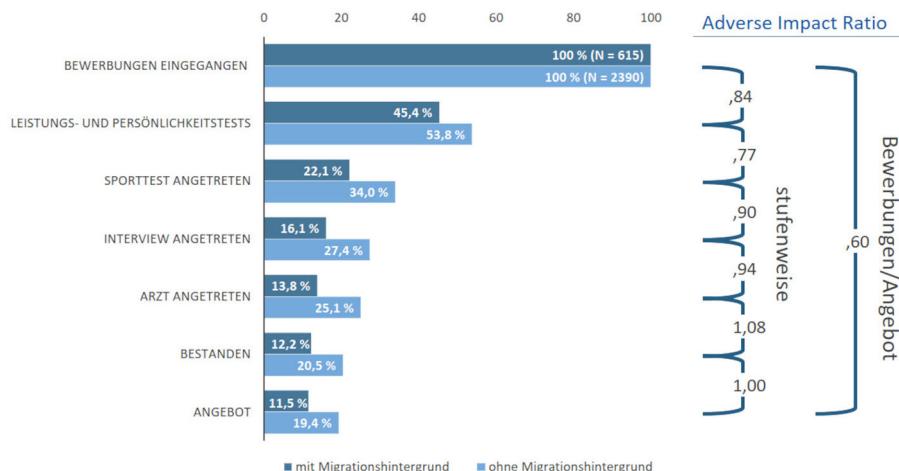
Der Vorteil von Cohen's d zur Beschreibung der Unterschiede zwischen zwei Gruppen ist wiederum, dass der Wert »unabhängig von willkürlich festgesetzten Selektionsquoten« (ebd.: 4) bewertet werden kann. Das heißt, Cohen's d steht für sich, indem es die Effektstärke des Unterschieds angibt und bewertbar macht und keinen Bezug etwa zur 4/5-Regel braucht. Nach der Konvention gilt ein d ab 0,2 als kleine Effektstärke, ab 0,5 als mittlere, und ab 0,8 als große Effektstärke. Ein positives Vorzeichen gibt ein besseres Abschneiden der Majoritätsgruppe an, ein negatives ein besseres Abschneiden der Minoritätsgruppe. Im Zusammenhang einer Varianzanalyse wird so beschreibbar, ob die charakteristischen Unterschiede innerhalb oder zwischen den Vergleichsgruppen liegen.

3.1 Standardisierte Selektion, blinde Benachteiligung

Eine direkte, merkmalsbezogene Diskriminierung, wie sie etwa aus der Arbeitsmarktforschung bekannt ist, bei der bei gleichen Bewerbungsunterlagen ein nicht deutsch klingender Name oder ein Bewerbungsfoto mit Kopftuch einen Unterschied mit Blick auf eine Einladung zum Vorstellungsgespräch machen (vgl. Weichselbaumer 2020; Weichselbaumer/Schuster 2021), ist für die ersten beiden Auswahlschritte der polizeilichen Verfahren formal ausgeschlossen. Durch den Ablauf des Verfahrens und die Reihenfolge der einzelnen Tests bleiben Bewerber:innen relativ lange anonym gegenüber der Organisation. So sind die polizeilichen Auswahlverfahren bis zum persönlichen Bewerbungsgespräch ›blind‹ gegenüber diskriminierungsanfälligen Merkmalen wie Hautfarbe, Geschlecht, Religionszugehörigkeit oder den Vor- und Nachnamen. Überprüft werden die Formalkriterien und dies zum Teil auch automatisiert im Rahmen der Online-Bewerbung.

Diese ›blinde Selektion‹ filtert bereits mehr als zwei Drittel aller Bewerber:innen entweder bei der Vorauswahl im Rahmen der Online-Bewerbung oder durch den kognitiven Leistungstest am PC heraus. Dabei sortiert letzterer Bewerber:innen mit Migrationshintergrund anteilig stärker heraus als etwa das persönliche Auswahlgespräch oder die anderen Testelemente (vgl. Abb. 1).

Abb. 1: Erfolgsquoten und anteiliges Chancenverhältnis nach Testelementen



In Abbildung 1 sind auf der linken Seite die Erfolgsquoten abgebildet; mit Referenz zur Anzahl der Bewerbungen für die jeweilige Gruppe ist angegeben, wie viel Prozent der Ausgangsgruppe die nächste Stufe im Verfahren erreichen. Von insgesamt 3005 Bewerber:innen gab 20,5 Prozent im Rahmen der Online-Bewerbung einen Migrationshintergrund an.¹⁹ Weniger als die Hälfte (45,4 Prozent) von diesen nahm am anschließenden Auswahlverfahren teil. Nach dem PC-Test verblieben noch 22,1 Prozent der migrantischen Bewerber:innen im Verfahren. Am Ende bestanden 12,2 Prozent der Bewerber:innen mit und 20,5 Prozent der ohne Migrationshintergrund das Verfahren. Ein Angebot für das Polizeistudium erhielten schließlich knapp jede:r fünfte Bewerber:in ohne Migrationshintergrund und etwa jede:r neunte migrantische Bewerber:in.

Das anteilige Chancenverhältnis der beiden Vergleichsgruppen liegt bei 0,6 zu 1. Dies ist auf der rechten Seite in Abb. 1 abgebildet. In Bezug auf das gesamte Verfahren liegt also ein Chancenverhältnis von 3/5 vor und damit ein *adverse impact*. Durch die stufenweisen Angaben des *adverse impact* zeigt die Abbildung außerdem, welche Testelemente überproportional zugunsten migrantischer Bewerber:innen gehen. Der PC-Test kann als die Hürde identifiziert werden, die die Chancen für migrantische Bewerber:innen besonders beeinträchtigt: Bei der Auswahl für den gehobenen Dienst in Rheinland-Pfalz selektiert der kognitive Leistungstest mit einem *adverse*

19 Rheinland-Pfalz ergebt den Migrationshintergrund als freiwillige Selbstangabe im Rahmen der Online-Bewerbung. Dabei ist die Definition des statistischen Bundesamtes hinterlegt und die Bewerber:innen können angeben, ob die Definition auf sie zutrifft.

impact von 0,77 (vgl. Abb. 1). Bei den anderen Stufen des Verfahrens liegen die Selektionsraten über 0,8 und deuten damit nicht auf gruppenspezifische Verzerrungen hin. Für die schrittweise Selektion wurde jeweils die Anzahl der Kandidat:innen berücksichtigt, die es bis zu dieser Teststufe geschafft hatte, so dass die Vergleichsgruppen pro Stufe kleiner und gewissermaßen ‚besser‘ werden. Damit bilden die Werte zwar die Selektionsquoten eines jeden Testelements ab, unterliegen aber einem ‚survivors‘ bias.

Ähnlich verhält es sich bei den Berechnungen für die Kohorte N₂ des Seiteneinstiegs in Rheinland-Pfalz: Die Erfolgschance migrantischer Bewerber:innen ist nur etwa halb so hoch (*adverse impact*: 0,55) wie bei ihren Mitbewerber:innen (ohne Abb.). Auch hier produziert der kognitive Leistungstest einen *adverse impact* von 0,51 und damit eine überproportionale Selektion migrantischer Bewerber:innen.

Für Bewerber:innen bei der Berliner Polizei lässt sich das Chancenverhältnis der hier betrachteten Vergleichsgruppen nach Sparten (Kriminalpolizei/Schutzpolizei) und Laufbahnen (gehobener Dienst, mittlerer Dienst) unterteilen, und zwar ist eine Bewerbung möglich für den gehobenen Dienst der Kriminalpolizei (gD Kripo), den gehobenen Dienst der Schutzpolizei (gD Schupo) und dem mittleren Dienst der Schutzpolizei (mD Schupo).

Abb. 2: Chancenverhältnis von Bewerber:innen mit und ohne Migrationshintergrund nach Laufbahnen über Zeit

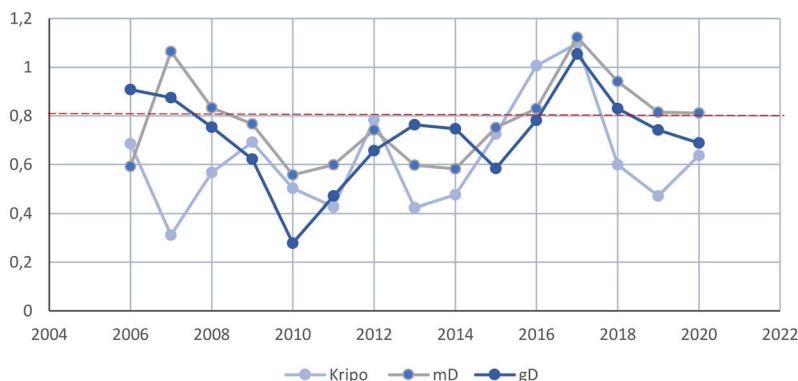


Abbildung 2 stellt die Entwicklung der Erfolgschancen von Bewerber:innen mit Migrationshintergrund für eine Einstellung bei der Berliner Polizei im zeitlichen Verlauf (2006–2020) dar. Differenziert nach Laufbahnen lässt sich erkennen, dass die Erfolgschancen migrantischer Bewerber:innen nicht stabil gleich verteilt sind. In den Jahren 2007 und 2008, sowie seit 2016 rekrutiert die Berliner Polizei im mitt-

leren Dienst chancengleich. Bei den Einstellungen im gehobenen Dienst gibt es für beide Sparten, Schutzpolizei und Kriminalpolizei, deutliche Schwankungen bzgl. der Erfolgschancen migrantischer Bewerber:innen.

Tab. 1: Korrelationsmatrix, Migrationshintergrund und Bestehen einzelner Tests, Berlin, Herbst 2020

	gD Kripo		gD Schupo		mD Schupo	
	N	Phi	N	Phi	N	Phi
Geschlecht	1632	-0,060*	1063	-0,102**	3331	-0,050**
Vortest	1321	-0,076**	878	-0,045	2246	-0,073**
PC-Test	772	-0,165**	568	-0,189**	1330	-0,123**
Sporttest	301	0,007	342	-0,007	547	-0,070
Personliche Vorstellung	192	-0,124	305	-0,099	404	-0,053
Leumundsprüfung	180	-0,057	298	-0,139*	407	-0,057
Tauglichkeitsprüfung	154	0,032	290	-0,053	341	-0,006

* Die Korrelation ist auf dem Niveau von 0,005 (2-seitig) signifikant.
 ** Die Korrelation ist auf dem Niveau von 0,001 (2-seitig) signifikant.
 Kodierung der Variablen: Migrationshintergrund (1=ja, 0=nein), Geschlecht (1=weiblich, 0=männlich), Tests bestanden (1=ja, 0=nein)

Für die Berliner Stichprobe zeigt sich, dass es kaum einen signifikanten Zusammenhang zwischen der Variable Migrationshintergrund und dem Bestehen verschiedener Testelemente gibt (vgl. Tab. 1). Die meisten Phi-Koeffizienten²⁰ liegen nahe Null und sind nicht signifikant, so beim Sporttest, der persönlichen Vorstellung, der Tauglichkeitsprüfung und der Leumundsprüfung für die Kripo und den mittleren Dienst. Allerdings scheitern Personen mit Migrationshintergrund, die sich für den gehobenen Dienst der Schutzpolizei bewerben, signifikant häufiger an der Prüfung des Leumunds ($\Phi=-0,139$). Hier ergeben sich Anschlussfragen für die Forschung, insbesondere danach, ob neben dem Führungszeugnis weitere Informationsquellen zur Feststellung des Leumunds, also des ›guten Rufes‹

20 Der Phi-Koeffizient ist ein Maß für die Stärke des Zusammenhangs zwischen zwei binären Variablen. Er kann auch als Maß für die Häufigkeit interpretiert werden, da er auf der Kreuztabelle der Häufigkeiten basiert. Ein Phi-Wert von 0 bedeutet, dass es keinen Zusammenhang gibt, während ein Wert von 1 oder -1 einen perfekten positiven oder negativen Zusammenhang anzeigt.

und der ›Unbescholtenheit‹, herangezogen werden, und welcher Bewertung sie unterliegen.²¹

Ein signifikanter negativer Zusammenhang mit der Variable Migrationshintergrund liegt außerdem für das Bestehen des Vortests und des PC-Tests vor, wobei der Zusammenhang für den PC-Test im gehobenen Dienst besonders negativ für migrantische Bewerber:innen ausfällt. Insgesamt besteht ein signifikanter Zusammenhang zwischen dem Bestehen des PC-Tests und der Variable Migrationshintergrund für alle Sparten und Laufbahnen. Das heißt auch: Die statistisch auffällige Selektion, die Anlass für eine Diskriminierungsvermutung ist, tritt beim PC-Test auf; obwohl dieser Test standardisiert ist und ohne Interaktion bewertet.

3.2 Verzerrungen innerhalb des standardisierten PC-Tests

In Gesprächen mit Personalverantwortlichen zu diesem Befund wurde *überwiegend* angenommen, dass Bewerber:innen mit Migrationshintergrund mit geringeren Bildungsvoraussetzungen starten und aufgrund geringerer Deutschkenntnisse schlechter in den PC-Tests abschneiden. Allerdings haben die Bewerber:innen mit Migrationshintergrund in den untersuchten Kohorten einen deutschen Bildungsweg und oft das deutsche Abitur. Der Verweis auf geringere Deutschkenntnisse schien daher nicht unmittelbar plausibel. Denn die an einer deutschen Schule erworbene Hochschulreife sollte auch die für die Polizei notwendigen Deutschkenntnisse umfassen.

Um besser zu verstehen, welche Aufgaben des PC-Tests migrantische Bewerber:innen stärker herausfordern, wurden die erzielten Punktewerte (*test scores*) einzelner Testteile für die Bewerbergruppen verglichen. Der kognitive Leistungs- test enthält mehrere einzeln bewertete Testelemente. Neben dem allgemeinen Intelligenztest werden etwa Sprachfähigkeit in Deutsch durch Rechtschreibung, Orthografie und Grammatik, allgemeine Merkfähigkeit, logisches, prozesshaftes und räumliches Denken sowie das Einschätzen beruflicher Situationen überprüft. Der folgende Vergleich bezieht sich auf die Testelemente, bei denen es um Deutschkenntnisse geht.

21 Ausgeschlossen werden sollte, dass Jugendliche, die von der Polizei häufiger kontrolliert wurden, häufiger in den polizeilichen Datenbanken auftauchen und deshalb einen negativen Leumund erhalten.

Tab. 2: Mittelwertdiskrepanz* für ausgewählte Testelemente und letzte Deutschnote

Auswahl (gD, 2020) Bachelor Polizeidienst RhPf.	Mig (nein)		Mig (ja)		Effektstärke	
	M	SD	M	SD	SD _{pooled}	D
Schulische Deutschnote (1276/279) ^{**}	7.9	2.3	7.1	2.2	13.5	-.05
<u>Kognitiver Fähigkeiten Test</u> (1286/279)						
Orthografie	50.4	7.3	48.0	8.0	7.5	-.32
Grammatik	50.6	7.2	47.2	8.3	7.5	-.45
Analogien	50.5	7.4	47.8	7.7	7.5	-.35
Situational Judgement Test	50.7	7.2	46.9	8.2	7.5	-.51
Kognitiver Fähigkeiten Test total	50.3	3.9	48.4	3.9	4.0	-.48

* Die Mittelwertdiskrepanz ist für alle hier angegebenen Werte auf dem Niveau von 0,001 signifikant.
** Die Differenz zwischen den 1286 Personen, die am kognitiven Fähigkeitentest teilgenommen haben, und der Anzahl der Bewerber:innen, für die eine schulische Deutschnote vorliegt, erklärt sich dadurch, dass bei Bewerber:innen mit Studium, die schulische Deutschnote nicht immer vorliegt und nicht nachgefordert werden muss.

Der Vergleich der Mittelwerte in Tabelle 2 zeigt einen signifikanten Unterschied der beiden Gruppen im Abschneiden bei den einzelnen Testelementen. In allen getesteten Bereichen des PC-Tests (Orthografie, Grammatik, Analogien, *Situational Judgement Test*) schneiden Bewerber ohne Migrationshintergrund besser ab. Die Effektstärken (d) sind negativ und variieren zwischen -0.32 und -0.51, was auf moderate Unterschiede hinweist. Besonders der *Situational Judgement Test* zeigt mit d = -0.51 die stärkste Differenz in den Leistungsunterschieden der beiden Bewerbergruppen an. Der Gesamtwert des kognitiven Fähigkeitstests zeigt ebenfalls eine moderate Effektstärke von d = -0.48.

Deutschkenntnisse: Bei den Noten im Abitur, die zwischen 0 und der Bestnote von 15 Punkten liegen können, ist der durchschnittliche Notenunterschied unter einem Punkt, Bewerber:innen ohne Migrationshintergrund haben im Durchschnitt eine um 0,8 bessere Punktzahl in der schulischen Deutschnote ($M_1 = 7.9$) als Bewerber:innen mit Migrationshintergrund ($M_2 = 7.1$). Die Effektstärke (d = -0.05) ist sehr gering, was darauf hindeutet, dass der Unterschied in der Deutschnote zwischen den beiden Gruppen minimal ist. Wird die Differenz zwischen den Deutschnoten im Abitur (d = -0,05) und dem Abschneiden in den polizeilichen Deutschtests (zwischen d = -0,32, und d = -0,45) verglichen, zeigt sich, dass die Leistungen der Gruppen in den polizeilichen Tests stärker auseinanderfallen als die Differenz in

den Schulnoten hätte vermuten lassen. Zusammenfassend lässt sich sagen, dass Bewerber ohne Migrationshintergrund tendenziell bessere Ergebnisse in den kognitiven Tests erzielen, die sich auf die Deutschkenntnisse beziehen, die mitgebrachten Deutschkenntnisse aber keinen signifikanten Unterschied aufweisen. Diese Ergebnisse widersprechen also der Annahme, dass schlechtere Bildungsvoraussetzungen migrantischer Bewerber:innen das schlechtere Abschneiden erklären. Seitens der Behörden wird argumentiert, dass die schulischen Deutschnoten auch weiche Faktoren wie Engagement berücksichtigen und deshalb besser ausfallen, da sie nicht nur Deutschkenntnisse abbilden. Gleichzeitig stellt sich die Frage, ob die Deutschkenntnisse des polizeilichen Tests in der Art berufsrelevant sind, dass sie die Selektion aus diesem Grund rechtfertigen.

Polizeiliche Situationen: Die stärkste Mittelwertdiskrepanz zwischen den Vergleichsgruppen resultiert aus dem *Situational Judgement Test* ($d = -0,51$). In diesem Test werden typische oder kritische Situationen aus dem Polizeialtag beschrieben. Per *multiple-choice* sollen die Kandidat:innen angeben, wie sie sich in den beschriebenen Situationen – etwa beim Überbringen einer Todesnachricht, oder bei Bedrohung mit einem Messer – als Polizist:in verhalten würden. Die starke Verzerrung zulasten migrantischer Bewerber:innen kann auf eine Kulturlastigkeit (cultural load) sowie eine Sprachlastigkeit (verbal load) des Tests zurückzuführen sein. Insbesondere die Kulturlastigkeit birgt hier ein Diskriminierungsrisiko, denn das Wissen über adäquates Rollenverhalten von Polizist:innen ist für Bewerber:innen vor der Ausbildung oder dem Studium eine Frage der Sozialisation. Für den Berufserfolg relevant ist aber, ob die professionelle Anforderung in der Ausbildung verinnerlicht und im Berufsalltag umgesetzt wird. Die Sprachlastigkeit des Tests meint, dass relativ viel Text gelesen und verarbeitet werden muss, obwohl das Testelement nicht die Sprachkompetenz überprüft. Beide Aspekte können zu der starken Differenz des *Situational Judgement Test* ($d = -0,51$) beitragen.

Bei entsprechenden Testelementen für die Laufbahnen in Berlin zeigt sich zwar auch eine Differenz zwischen Bewerber:innen mit und ohne Migrationshintergrund, ihr Effekt ist aber schwach: Für die Testelemente »Intelligenz«, »Polizeiliche Situationen«, »Deutschtest«, »Politik und Gesellschaft« liegen die d -Werte zwischen $d = -0,1$ (Politik und Gesellschaft für mD Schupo) und $d = -0,32$ (Deutschtest gD Schupo). In Berlin unterscheiden sich Bewerber:innen mit und ohne Migrationshintergrund für den gehobenen Dienst in den schulischen Deutschnoten stärker als im polizeilichen Deutschtest. Hier verhält es sich also genau umgekehrt zur Bewerberkohorte für den gehobenen Dienst in Rheinland-Pfalz. Bei den Bewerbungen für den mittleren Dienst in Berlin gibt es keinen Zusammenhang zwischen dem Abschneiden von Bewerber:innen und der Variable »Migrationshintergrund«.

Mit Verweis auf Berliner Polizei wird häufig angeführt, dass sie so repräsentativ rekrutiere, weil es in Berlin noch den mittleren Dienst gebe. Das schließt die Annahme ein, dass migrantische Bewerber:innen auf geringere Bildungsvoraussetzungen

angewiesen sind, um zur Polizei zugelassen zu werden. Ein Vergleich von Abiturient:innen bzgl. der angestrebten Laufbahn bei Bewerbung zur realisierten Laufbahn bei Einstellung zeigt allerdings, dass diese Vermutung zu einfach ist. Denn Abiturient:innen mit Migrationshintergrund bleiben überdurchschnittlich häufig hinter den Möglichkeiten ihres Bildungsabschlusses zurück (vgl. Tab. 3).

Tab 3: Laufbahnaspirationen und -einstellungen von Abiturient:innen, Berlin, Jahrgänge 2011 bis 2020

Laufbahn bei Bewerbung	Mig.	Laufbahn bei Einstellung			
		gD Kripo	gD Schupo	mD Schupo	Total (N)
gD Kripo	Ohne	65,7 %	33,0 %	1,3 %	100,0 %
					(1657)
	Mit	57,3 %	39,6 %	3,0 %	100,0 %
					(497)
gD Schupo	Ohne		70,2 %	29,8 %	100,0 %
					(2324)
	Mit		65,6 %	34,4 %	100,0 %
					(834)
mD Schupo	Ohne			100,0 %	100,0 %
					(900)
	Mit			100,0 %	100,0 %
					(412)
				Gesamt N	6624

Die Tabelle zeigt, für welche Laufbahn und Sparte der Berliner Polizei sich Abiturient:innen zwischen 2011 und 2020 beworben haben, und ob sich der Wunsch realisiert hat. Aufgeführt sind nur eingestellte Personen, die sich mit Abitur beworben haben und damit für den gehobenen Dienst qualifizieren.

Zunächst fällt auf, dass das Verhältnis der absoluten Bewerbungen sich laufbahnbezogen unterscheidet: Abiturient:innen mit Migrationshintergrund bewerben sich anteilig häufiger für den mittleren Dienst als für den gehobenen Dienst: Sie stellen 31,4 Prozent der Bewerbungen für den mittleren Dienst, 26,4 Prozent der Bewerbungen für den gehobenen Dienst der Schupo und 23,1 Prozent der Bewerbungen für die Berliner Kriminalpolizei.

1657 Personen mit Abitur und ohne Migrationshintergrund haben den gehobenen Dienst bei der Kripo angestrebt und das Auswahlverfahren bestanden. Von diesen sind 65,7 Prozent tatsächlich auf die gewünschte Stelle eingestellt worden. Von 497 Personen mit Abitur und Migrationshintergrund waren es hingegen 57,3 Prozent. Für 1,3 Prozent der nicht-migrantischen Bewerber:innen und 3 Prozent der migrantischen Bewerber:innen konnte die Kripo nicht realisiert werden, so dass sie ihre Laufbahn im mittleren Dienst aufnahmen. Von den Bewerbungen für den gehobenen Dienst der Schutzpolizei erreichten 29,8 Prozent ohne Migrationshintergrund und 34,4 Prozent mit Migrationshintergrund den mittleren Dienst. Bewerber:innen mit Migrationshintergrund bleiben also anteilig häufiger hinter den Möglichkeiten ihres Bildungsabschlusses zurück als die Vergleichsgruppe. Schließlich haben wir für die Kohorten aus Rheinland-Pfalz in einer hierarchischen Regressionsanalyse geprüft, ob die Variable Migrationshintergrund ein eigener Erklärungsfaktor für das Abschneiden in den Testverfahren ist. Dies lässt sich für die Bewerbergruppen aus Rheinland-Pfalz für den PC-Test und das Interview bestätigen, auf die Variable Migrationshintergrund entfällt unter Kontrolle der Variablen Geschlecht, Alter und Bildung noch 4,1 Prozent der Varianzaufklärung. Dies entspricht den Ergebnissen für NRW von Stumpf et al. (2016: 19, 23), nach denen die Variable Migrationshintergrund unter Kontrolle noch 4,5 Prozent der Varianzaufklärung ausmachte. Ihre Schlussfolgerung kann folglich unterstrichen werden: ein Verweis auf geringere Bildungsqualifikationen reicht nicht aus, um das schlechtere Abschneiden von Bewerber:innen mit Migrationshintergrund zu erklären.

4. Alchemische Bewertungen und organisationskulturelle Hürden

Ich komme nun zum kommunikativen Teil des Auswahlverfahrens, zur sogenannten interaktiven Diagnostik. Dieser Teil findet in den meisten Bundesländern in Form eines persönlichen Auswahlgesprächs oder klassischen Jobinterviews statt. Diesem liegt in der Regel ein standardisierter Fragenkatalog zugrunde. Einige Polizeien setzen auch Referate, Rollenspiele oder Gruppendiskussionen ein und weisen den kommunikativen Teil des Auswahlverfahrens als Assessment-Center aus. Hier werden Berufsmotivation, Organisationswissen, Stress- und Konfliktbewältigung, Selbstreflexion, politische Bildung, Ausdrucksfähigkeit und das Auftreten beurteilt. Die Beurteilung erfolgt in Punktwerten, die für einzelne Kompetenzen oder den gesamten mündlichen Teil vergeben werden. Die Bepunktung wird von den Bewertenden oder einer Auswahlkommission direkt im Anschluss an die mündliche Testsituation vorgenommen. Die Bewertenden, polizeiintern auch als *rater* bezeichnet, sind zum einen Polizist:innen, die in dieser Funktion teils haupt- und teils nebenamtlich tätig sind, oder zivile Angestellte der Ausbildungseinrichtungen oder Hochschulen wie Lehrbeauftragte für Deutsch oder Politikwissenschaft. Die *rater* werden

auf die Personenbewertung durch Schulung vorbereitet. Die Bewertungsteams waren in den beobachteten Auswahlsituationen paritätisch besetzt. Einen Migrationshintergrund nannte keine:r der Bewertenden.

Im Fokus der Untersuchung stehen nun die Bewertungen der *rater* bzw. die Begründungen, die sie formuliert haben, um ihre Bewertung der Kandidat:innen zu erklären, herzuleiten und zu rechtfertigen. In Anlehnung an Gomolla und Radtke (2009: 152) gehe ich davon aus, dass die Bewertenden in ihren Argumentationen eine institutionelle Wirklichkeitskonstruktion bekräftigen. Das heißt, dass sie

»genau die Deutungsmuster zur Plausibilisierung und Legitimierung ihrer Entscheidung [hier: Bewertung] anführen, von denen sie sicher wissen, dass sie institutionell approbiert sind und als Kommentierung der eigenen Praxis akzeptiert werden.« (Ebd.)

Die Bewertenden ordnen mit ihren Ausführungen das eigene Urteil »in den Bereich des Gültigen« (ebd.) ein. Dies tun sie regulär gegenüber den Mitbewertenden, in den beobachteten Situationen galten die Rechtfertigungen aber auch uns teilnehmenden Beobachterinnen. Im Fokus stehen Argumentationen²², in denen die Differenz »Migrationshintergrund« nicht nur angemerkt wird, sondern als Deutungsressource in die Bewertung einfließt.

Ausgehend von einem Beispiel werden in den nächsten beiden Abschnitten Bewertungen zur ›Interkulturellen Kompetenz‹ (4.1) sowie zu Berufsmotivation und Werdegang (4.2) dargestellt und analysiert. Das Beispiel ist explizit ein Negativbeispiel. Es steht nicht für die Qualität der mündlichen Testsituationen und kann nicht generalisiert werden. Wohl aber reicht es, um auf Diskriminierungsrisiken aufmerksam zu machen. In dem Bewertungsvorgang kommt nahezu prototypisch zum Ausdruck, was Robert Merton (1948) als Moral-Alchemie beschrieben hat: die »engagingly simple formula«, nach der »the same behavior must be differently evaluated according to the person who exhibits it« (ebd.: 201). Dadurch kommt es zu paradoxen, zu alchemischen Bewertungsvorgängen. Den folgenden zwei Abschnitten ist eine Sequenz aus einem Beobachtungsprotokoll vorangestellt, die eine zusammenhängende Bewertung bilden. Die zitierten Aussagen wurden in genau dieser Reihenfolge getätigt. Als dritten Punkt ergänze ich die Bewertung von Deutsch- und zusätzlichen Sprachkenntnissen im Rahmen der Auswahlverfahren, da auch an diesem Beispiel eine institutionelle Bewertung deutlich wird, die mit zweierlei Maß misst (4.3).

22 Im Modus der Argumentation werden empirische und normative Behauptungen vertreten, die mit dem Anspruch auf Gültigkeit vorgetragen werden. Dabei sind Argumente im Kontext Organisation zu unterscheiden vom Modus der Verhandlungen, in dem Forderungen und Kompromisse formuliert werden (vgl. Gomolla/Radtke 2009: 152).

4.1 Bewertung von ›Interkultureller Kompetenz²³

»Irgendwie kann ich ihn nicht so recht greifen. (...) Was mich total enttäuscht hat, war die Interkulturelle Kompetenz. Also, wenn jemand selber Migrationshintergrund hat, da hätte ich mehr erwartet. Da hätte ich erwartet, dass der einfach so blubbert, ›Das kenn ich von zu Hause, da ziehen wir auch immer die Schuhe aus, oder so was. Da so oberflächlich zu bleiben, (...) find ich, (...) naja. Da sehe ich ihn unterdurchschnittlich in der Interkulturellen Kompetenz.«²⁴

Die protokollierte Situation gibt die Besprechung zwischen einem hauptamtlichen Bewerter und einer nebenamtlichen Bewerterin zu einem 45-minütigen Auswahlgespräch mit einem Kandidaten wieder, dem von beiden Bewertenden ein Migrationshintergrund attestiert wird. Der Interviewte hatte den Raum bereits verlassen. Die Bewerterin eröffnet das Gespräch und umschreibt zunächst ihre Unsicherheit bei der Bewertung des Kandidaten, den sie »irgendwie ... nicht so recht greifen« könne. Als Anhaltspunkt für die Bewertung geht sie dann als erstes auf die ›Interkulturelle Kompetenz‹ ein, die sie »total enttäuscht« habe. Die Differenzierung Migrationshintergrund wird genutzt, um den Kandidaten greifbar zu machen. Gleichzeitig wird an diese Differenzierung als Deutungsressource angeschlossen: »Wenn jemand selber Migrationshintergrund hat, da hätte ich mehr erwartet«. Die Bewertungsunsicherheit resultiert demnach aus einer Diskrepanz zwischen Wahrnehmung und Erwartung. Letztere wird von der Sprecherin dann auch konkretisiert: Als Erwartung spielt sie auf ein Narrativ an, das polizeiintern regelmäßig als Beispiel für interkulturelle Konfliktsituationen genannt wird: beim Einsatz in muslimischen Haushalten würden Polizist:innen dazu aufgefordert, ihre Schuhe auszuziehen, was die Eigensicherung gefährde.²⁵ Dieses polizeiliche Narrativ wird in der Bewertungssituation zur professionellen Anforderung. Der Kandidat sollte nicht nur mit diesem polizeiinternen Narrativ vertraut sein, sondern sich auch offenbaren und berichten, dass er es »von zu Hause« nicht anders gewohnt sei, weil »da ziehen wir auch immer die Schuhe aus«. Eingefordert wird also einmal, dass der Kandidat die Perspektive der Polizei übernimmt, indem er das

23 Für eine kritische Auseinandersetzung mit der Operationalisierung von Interkultureller Kompetenz in den polizeilichen Personalauswahlverfahren vgl. Leenen et al. (2014).

24 Zögerliches Sprechen oder längere Pausen sind mit drei Punkten in runden Klammern (...) notiert.

25 Polizeiintern gilt dies als Beispiel für die Spannung zwischen religiöser Autorität und staatlicher Autorität. Teilweise würden Muslime die religiöse Autorität über die staatliche Autorität stellen und damit auch von Polizist:innen im Einsatz verlangen, die Schuhe auszuziehen. Respekt gegenüber Kultur oder Religion dürfe nicht auf Kosten der Eigensicherung gehen. In einigen Beispielen ist es nur der Teppich der nicht mit Schuhen betreten werden darf und nicht die Wohnung.

Ausziehen der Schuhe in bestimmten Haushalten als Beispiel für interkulturelle Konflikte benennen kann, gleichzeitig wird von ihm erwartet, die Verhaltensweise des polizeilichen Gegenübers zu authentifizieren. Als migrantisch markierte Person soll er die Realitätstreue des Beispiels bestätigen, gar belegen. So wird ein in der Prüfung kontextualisiertes *Othering* zur paradoxen Erwartung, die zudem eine stereotypisierende Problembeschreibung legitimiert.

Da der Kandidat nicht wie erwartet antwortet, wird er als »unterdurchschnittlich« in der Interkulturellen Kompetenz bewertet, seine Antwort als »oberflächlich« empfunden. Die Enttäuschung gilt also paradoxe Weise der Tatsache, dass er dem Klischee der Bewerterin nicht entsprochen hat und nicht einfach so »blubbert«, als es darum geht, »von zu Hause« zu erzählen und dadurch die unterstellte Interkulturelle Kompetenz unter Beweis zu stellen. Die enttäuschte Erwartung führt dazu, an der Authentizität des Bewerbers zu zweifeln, er wirkt nicht greifbar. Bereits der Einstieg in die Bewertungssituation transportiert einen latenten Zweifel, der sich im weiteren Verlauf auch auf die Bewertung der Berufsmotivation und des Bildungswegs auswirkt.

Da der migrantisch gelesene Bewerber im Verlauf des Vorstellungsgesprächs mit seinen Antworten nicht die Erwartungen der Bewertenden erfüllt, wurde ihm zusätzlich eine weitere Frage zum Themenbereich Interkulturelle Kompetenz gestellt, die den meisten anderen Bewerber:innen nicht gestellt wurde. Er sollte sich vorstellen, dass er mit einem Freund einer bestimmten Nationalität auf ein Konzert gehen wolle. In der Rahmung der Fragestellung sollte diese Nationalität für Unpünktlichkeit stehen.²⁶ Die Frage war nun, wie der Kandidat seinen Freund zu Pünktlichkeit veranlassen würde. In der Problembearbeitung des Kandidaten bleibt die Nationalität jedoch unerwähnt, wie dem Beobachtungsprotokoll zu entnehmen ist:

Erschildert, dass er mehrfach im Laufe des Tages vorher mit dem Freund in Kontakt treten würde; und ihn dann vielleicht einfach direkt von zu Hause abholen würde, um zusammen zum Konzert zu gehen. Den Bewertern fehlt an dieser Stelle scheinbar die Kulturalisierung des Problems; sie haken etwas suggestiv nach: »Woran könnte es denn liegen, dass Ihr Freund immer zu spät kommt?« Der Bewerber ist verunsichert, antwortet wieder hastend, fragend: »Vielleicht ist ihm das nicht so wichtig? Vielleicht versteht er meine Sprache nicht?« Daraufhin wird nochmal nachgehakt: »Wie ist es denn bei Ihnen, sind Sie immer pünktlich?« »Ja«, antwortet der Bewerter prompt und fast mit soldatischem Ton. »Ich bin immer pünktlich. Ich bereite mich vor auf das, was ich vor hab.«

26 Dabei wichen die *rater* vom Skript ab und nutzten zur Charakterisierung eine andere Nationalität als die, die der standardisierte Fragebogen vorsah.

Der Kandidat wollte sich also nicht der stereotypen Idee über unpünktliche Nationalitäten anpassen, wohl aber dem Klischee deutscher Pünktlichkeit entsprechen. Das vorsichtige, tastende Fragen wird als Mangel interpretiert, da der Kandidat das Problem scheinbar nicht benennen kann. Als Beobachterin dieser Situation fiel mir auf, dass scheinbar nicht die Lösungsstrategie bewertet wurde, sondern die fehlende Problembeschreibung.

Die Prüfsituation erschien mir derart absurd, dass ich meinen Eindruck später im Gespräch mit den für das Verfahren zuständigen Expert:innen thematisierte. Dabei wurde deutlich, dass Interkulturelle Kompetenz ein umstrittenes Prüfkriterium ist. Eigentlich sei es keine klassische Kompetenz, sondern eher eine Liste von Fähigkeiten und Einstellungen, darunter Offenheit und Ambiguitätstoleranz, die sowieso schon geprüft würden. Interkulturelle Kompetenz als eigenes Prüfkriterium heranzuziehen, wurde unter anderem nach dem Auftauchen rechter Chatgruppen politisch vorgegeben, »damit man da was nach außen hat«. Deutlich wird ein irreführender Zusammenhang zwischen Extremismusprävention und der Konzeption Interkultureller Kompetenz. Dieser Konnex ist auch unter den Verantwortlichen umstritten. In der Anwendung ergeben sich Diskriminierungsrisiken für migrantische Bewerber:innen, indem Stereotypen nicht nur Raum zur Entfaltung bekommen, sondern auch als Anforderung gelten.²⁷

4.2 Bewertung von Berufsmotivation und Werdegang

Es wird noch über den Wohnort des Bewerbers gesprochen, und dass es ja nicht sein könnte, wenn man dort wohne, dass man keine Erlebnisse und Konflikte mit kulturellen Problemen schildern könne. Die Sicht des Bewerbers sei »viel zu optimistisch«, auch auf den Beruf.»Ich wusste nicht, ob er tatsächlich emotional gefestigt ist, oder das einfach nur noch nicht abschätzen kann. (...) Also die Selbsteinschätzung und Selbstwahrnehmung scheint mir auch nicht ganz zu passen.«

»Auch beim Konzeptionellen Arbeiten, das fand ich suboptimal, also ich weiß ja nicht wie gut die Schule war, oder das Umfeld, oder so.«

Im Anschluss an die Bewertung der Interkulturellen Kompetenz folgt eine Argumentation, in der der Wohnort als Deutungsressource herangezogen wird. Die Bewertenden tauschen sich über den Wohnort des Kandidaten aus, über den er selbst nur indirekt gesprochen hatte, indem er von »Freunde[n] aus unserem Viertel«, mit denen er »gut klarkomme«, sprach. Die Bewerterin verbindet mit dem Stadtteil offenbar Negatives und hält es für wenig plausibel, »wenn man dort wohne, dass man

27 Die Lösung ist das Problem – wie Fatoş Atali Timmer und Ikram Errahmouni Rimi (2024) es treffend formuliert haben.

keine Erlebnisse und Konflikte mit kulturellen Problemen schildern könne«. Mit Referenz auf den Wohnort wird das Antwortverhalten des Kandidaten erneut in Zweifel gezogen. Auf die Frage, ob er von Erfahrungen mit kulturellen Konflikten berichten könne, hatte er keine offenen Konflikte thematisiert, sondern beschrieben,

»dass man kulturelle Unterschiede möglicherweise dadurch erfahren kann, wenn man zu jemanden nach Hause gehe. Die Wohnungseinrichtung sei möglicherweise anders und auch das Essen und der Umgang mit den Eltern. Ein Freund von ihm sei nicht aus Deutschland, und in seiner Familie werde es als unhöflich angesehen, wenn man sich zu den Eltern ins Wohnzimmer setze, was bei vielen deutschen Familien aber normal sei.«

In der Beurteilung führt diese Antwort dazu, dass an seiner Wahrnehmungsfähigkeit und Selbsteinschätzung gezweifelt wird. Die dargestellten Wahrnehmungen zu kulturellen Unterschieden werden als »viel zu optimistisch« eingeschätzt, wie auch seine Sicht auf den Polizeiberuf. Daher fällt auch die Beurteilung seiner Berufsmotivation verhalten aus.

In anderen Auswahlgesprächen wurde Berufsmotivation positiv bewertet, wenn die Kandidat:innen das »Organigramm kannten«²⁸, viel Organisationswissen bewiesen und mögliche Arbeitsbelastungen nennen konnten. Diese Anforderung wurde zwar von dem Kandidaten erfüllt, indem er das Wissen über die Organisation wiedergab und körperliche und psychische Belastungen wie Nachschichten thematisierte. Allerdings wurde seine Zuversicht, damit umgehen zu können, nicht ernst genommen. Die Bewertenden glaubten nicht, dass seine Antwort auch tatsächlich auf ihn zutreffe. »Ich wusste nicht, ob er tatsächlich emotional gefestigt ist, oder das einfach nur noch nicht abschätzen kann.« Die umgekehrte Situation trifft auf Bewerber:innen zu, die auf Freunde oder Verwandte bei der Polizei verweisen, das waren 9 von 32 Kandidat:innen. Das Nennen von Freunden und Verwandten bei der Polizei wurde in diesen Gesprächen in den allermeisten Fällen als Berufsmotivation gedeutet. Auch unter den *rater* war dies ein gängiger Verweis, um Berufsmotivation zu dokumentieren, die Kandidat:innen wüssten besser, worauf sie sich einlassen und was das für das Familienleben bedeute. In einer Situation entschuldigen die Bewertenden den Bericht eines männlichen Kandidaten ohne Migrationshintergrund, der nach eigenen Angaben in der Ausbildung nicht mit den Kollegen klarekommen sei, mit einem Verweis auf seine Eltern: »Er hat ja auch ne barsche Art, kriegt er wahrscheinlich auch vorgelebt von seinen Eltern, typische Polizisten Art, typisch für unsere Generation Z.«

28 Die Phrase »Kann das Organigramm« wurde als Notation vereinbart, wenn Kandidat:innen den Aufbau der Organisation mit personeller Besetzung der Posten nennen konnten. Ge nannt wurden meistens der Innenminister als Dienstherr, der Polizeipräsident/die Polizeipräsidentin, sowie in Bundesländern die Anzahl der Direktionen.

Beide Bewertungen begründen eine Verletzung biografischer Fairness, denn nur selbst zu verantwortende biografische Indikatoren dürfen in der Personalauswahl Berücksichtigung finden – der Beruf der Eltern, der Wohnort der Familie und der Ruf der Schule gehören nicht dazu.

Zurück zum Ausgangsbeispiel: Der Kandidat hatte den Polizeiberuf als Hauptmotivation für seinen Bildungsweg vom Hauptschulabschluss zur Fachhochschulreife genannt. Der Bildungsaufstieg, der nach Angaben des Kandidaten durch den »Traumberuf Polizist« motiviert war, wurde im Gegensatz zur Nennung von Freunden und Verwandten bei der Polizei nicht als Berufsmotivation gewertet.

Bildungsvoraussetzungen, die als Teil der Formalkriterien bereits im ersten Teil des Bewerbungsverfahrens formal bewertet wurden, werden hier relativiert. Durch Referenz auf einen diskriminierenden Wissensbestand aus dem Bildungsbereich²⁹ wird eine Anforderung, die formal gleich zu behandeln war, nun doch ungleich bewertet und zwar erneut mit Bezug zum Wohnort. Wieder werden Zweifel formuliert: »also ich weiß ja nicht wie gut die Schule war, oder das Umfeld, oder so.«

Die Deutungsressource Wohnort – »wenn man da wohnt«, »oder das Umfeld oder so« – prägt die Bewertung des Kandidaten stark. Sie führt dazu, dass Interkulturelle Kompetenz, Berufsmotivation und selbst der Bildungsaufstieg nicht erkannt und die Antworten des Kandidaten hinterfragt werden. In der protokollierten Situation werden drei Verzerrungen deutlich, die aufgrund der zugeschriebenen sozialen Identität entstehen und in der Bewertung eine »Moral-Alchemie« nach sich ziehen. Dabei werden Personen, die vermeintlich einer *out-group* zugerechnet werden, für positive Eigenschaften (»Tugenden« bei Merton) negativer bewertet als Personen, die der *in-group* zugerechnet werden.

- Die Bewertung von Interkultureller Kompetenz führt zum *Othering*. Soziale Kompetenzen im Umgang mit Unterschieden können nicht zum Tragen kommen, da Stereotype nicht nur Raum zur Entfaltung bekommen, sondern als Anforderung gelten.
- Die Bewertung der Bildungsvoraussetzungen führt über eine Referenz zum Wohnort zur Verletzung biografischer Fairness. Der Werdegang und der schulische Aufstieg verwandelten sich durch den Wohnort in Zweifel an der Selbsteinschätzung des Bewerbers und Zweifel an der Ordentlichkeit oder Gleichwertigkeit seiner Schulbildung.

²⁹ Paul Mecheril stellte bspw. fest, dass im Bildungsbereich in Deutschland – ausgehend von einem homogenen Nationenbegriff – ständig zwischen Menschen mit und ohne »Migrationshintergrund« unterschieden werde. Hierbei würden Menschen mit Migrationsgeschichte bestimmte negative, von der Norm abweichende Eigenschaften zugeschrieben, die folglich problematisiert werden. Positive Eigenschaften dieser Menschen würden ausgeblendet (vgl. Mecheril 2015: 29–34).

- Die Bewertung der geäußerten Berufsmotivation führt zu Zweifel an der Selbst-einschätzung und emotionalen Reife des Bewerbers. Die an den Tag gelegte und in Bewerbungsgesprächen sozial erwünschte Berufsmotivation wurde nicht als Motivation, sondern als unreifer Optimismus gedeutet.

Insgesamt kam es in den beobachteten Auswahlgesprächen häufig zu stereotypisierenden Bewertungen, die sich insbesondere auf die Herkunft bezogen, sowohl auf die kulturelle als auch die soziale Herkunft. Allerdings führten diese Bewertungen in den beobachteten Fällen nicht zu einem Ausschluss der Kandidat:innen aus dem Verfahren. Vielmehr wurden diese Bewertungen durch die standardisierten Bewertungsbögen und die angeleitete Bepunktung abgemildert. Dadurch fielen stereotype Bewertungen in den beobachteten Situationen nur punktreduzierend ins Gewicht, was zwar nicht direkt zum Ausschluss führt, aber einen niedrigeren Ranglistenplatz nach sich ziehen kann. Eine direkte Diskriminierung wird also methodisch kontrolliert, eine mittelbare Diskriminierung entzieht sich jedoch ihrer Feststellung. Vorurteilsgeleitete Bewertungen werden zwar durch die Bepunktung abgemildert, in letzter Konsequenz aber auch ungreifbar.

4.3 Bewertung von Sprachkenntnissen

Die Alltagsrelevanz zusätzlicher Sprachkenntnisse für die Polizeiarbeit ist durch ihre fehlende institutionelle Anerkennung ein weiteres Beispiel für eine Bewertungs-alchemie. Zwar gelten Sprachkenntnisse als nützliche Ressource (insbesondere für die Polizeiorganisation), für die Sprachkompetenten selbst haben sie keinen Vorteil. Beim Zugang zum Polizeidienst spielten Sprachkenntnisse seit den 1990er Jahren insbesondere für in Deutschland lebende Ausländer eine Rolle. Mit Kenntnissen einer Bedarfssprache konnten sie nach der seit 1993 existierenden Ausnahmeregelung im Beamten gesetz Beamte des deutschen Staates werden. Als Adressaten von Art. 33 (2) sind Deutsche mit Migrationshintergrund jedoch nicht auf eine solche Ausnahme angewiesen. Sie gelangen regulär über die Auswahlverfahren in die Polizei. Im Unterschied zu anderen Organisationen mit viel Publikumskontakt werden zusätzliche Sprachkenntnisse im Auswahlverfahren kaum berücksichtigt. In den Auswahlverfahren der Polizeien von Bund und Ländern gibt es dazu keine einheitliche Verfahrensweise. In Rheinland-Pfalz etwa können zusätzliche Sprachen im Rahmen der Online-Bewerbung eingetragen werden und fallen als zusätzliche Punkte im Ranking der Vorauswahl mit ins Gewicht. In Berlin gab es für einige Jahre die Möglichkeit, für weitere Sprachen im Auswahlverfahren einen Test zu machen und zusätzlich zu punkten. Darauf wurde nach Auskunft der Behörde jedoch nur selten zurückgegriffen, so dass das Angebot wieder abgeschafft wurde.

Die Polizeien tun sich eher schwer, zusätzliche Sprachkenntnisse als berufs-relevante Fähigkeiten regulär und institutionell anzuerkennen. Im Austausch mit

Verantwortlichen der Einstellungsstellen wurden drei grundsätzliche Argumente gegen eine institutionalisierte Anerkennung zusätzlicher Sprachkenntnisse deutlich: 1) Unzulässige Kompensation schlechter Deutschkenntnisse, 2) Benachteiligung nicht-migrantischer Bewerber:innen und 3) Fehlende Prüfbarkeit des Sprachniveaus.

Erstens dominiert die Befürchtung, durch eine Anerkennung zusätzlicher Sprachkenntnisse würden die Ansprüche an die Deutschkenntnisse gesenkt. Meine Gesprächspartner:innen assoziierten Migrationshintergrund ausnahmslos mit schlechteren Deutschkenntnissen. Mit dieser stereotypen Grundannahme wurde ein (zusätzliches) Anrechnen von Sprachkenntnissen als (rechtlich) unzulässige Kompensation schlechter Deutschkenntnisse gedeutet. An diese Sorge vor einer Relativierung von Leistungen im Auswahlverfahren knüpft mit Blick auf das spätere Berufsleben die Sorge an, dass bürokratische Standards nicht erfüllt werden. Es wurde argumentiert, dass das, was ein Polizist schreibt, auch vor Gericht Bestand haben müsse. Polizisten müssten deshalb in der Lage sein, gerichtsfeste Akten zu produzieren. Interessant ist hier der fließende Übergang von einer Sorge der Relativierung von Leistung zu einer Sorge um das Funktionieren des Rechtsstaats.

Dazu äußerte sich eine Führungskraft in einer bundesweiten Diskussionsrunde: »Unsere Auswahlverfahren sind so gut, dass sie auch ohne Sonder-Bepunktung viele Bedarfssprachen in die Polizei holen.« Hier wird den Auswahlverfahren zugeschrieben, was den Bewerber:innen gebührt, und das obwohl die Verfahren eben nicht dafür sorgen, dass Sprachkenntnisse gefördert werden. Durch die Argumentation, die beides lobt – Auswahlverfahren und Ergebnis –, wird unterstrichen, dass es keinen Grund für strukturelle Anpassungen gibt.

Zweitens wird eine Bepunktung zusätzlicher Sprachkenntnisse als Benachteiligung nicht-migrantischer Bewerber:innen gesehen, da diese keine Möglichkeit hätten, zweisprachig aufzuwachsen. Durch die ungleiche Ausgangssituation würden Punkte vergeben, die nicht alle Kandidat:innen in gleicher Weise erlangen könnten. In dieser Argumentation wird der alchemische Charakter der Bewertung besonders deutlich, denn umgekehrt wird dieses Argument nicht vorgebracht: Kandidat:innen mit schlechten Deutschkenntnissen, die in der Regel eine Frage häuslicher Ressourcen sind, wird keine Benachteiligung zuerkannt.

Drittens sprechen ökonomische Gründe gegen eine institutionelle Anerkennung zusätzlicher Sprachkenntnisse. Eine Anerkennung im Rahmen der Auswahlverfahren gehe mit der Verpflichtung für die Einstellungsstellen einher, die jeweiligen Sprachen auch auf dem festgelegten Niveau zu überprüfen. Dadurch entstünden bei der Vielzahl an Sprachen hohe zusätzliche Personal- und Sachkosten. Skeptisch stehen die Befragten einem Beleg durch einen externen Sprachtest gegenüber, wie etwa an Universitäten üblich.

Es gibt auch einzelne Stimmen in der Polizei, die eine institutionelle Anerkennung zusätzlicher Sprachen befürworten. Eine Gleichstellungsbeauftragte setzte

sich etwa in einer Gruppendiskussion für eine bessere Anerkennung ein, indem sie auf einen Fall doppelter Benachteiligung hinwies: Ein Polizeianwärter hatte ihr berichtet, dass er im Praktikum 80 Prozent seiner Zeit mit Übersetzungstätigkeiten beauftragt worden war. Anschließend erhielt er eine negative Bewertung, da er zu wenig praktische Erfahrungen gesammelt habe. Die Polizistin sah hier eine Benachteiligung, da der Praktikant ja auf Anweisung übersetzt und zudem eine Arbeit geleistet habe, die sonst niemand hätte übernehmen können. Dass er dafür negativ bewertet wurde, betrachtet sie als Diskriminierung, für die sie in der Diskussion sensibilisieren wollte. In der Runde wurde diesem Bericht Unverständnis entgegengebracht, »das könne man sich beim besten Willen nicht vorstellen«. Auch wurde diskutiert, warum sich der Praktikant nicht beschwert habe.

In Interviews und Diskussionsrunden berichteten viele sprachkompetente Polizist:innen, dass sie neben den Übersetzungen im Dienst auch in der Freizeit angerufen werden, um Übersetzungsarbeit zu leisten. Auf Nachfrage gaben die meisten von ihnen an, mehrmals die Woche für Übersetzungen oder fremdsprachliche Vermittlungsarbeit angerufen zu werden. Auf erneute Nachfrage bezifferten sie solche Anfragen in der Freizeit mit drei- bis fünfmal pro Woche. Die Qualifikation wird also auf dem kurzen Dienstweg in einer Regelmäßigkeit genutzt, die über eine gelegentliche kollegiale Gefälligkeit hinausgeht. Alexandra Graevskaia (2022) markiert diesen strukturellen Bedarf an zusätzlichen Sprachen im Polizeialltag gar als »Unterentbehrlichkeit«. Behr und Molapisi (2022: 85) beschreiben in diesem Zusammenhang eine gewisse Unterordnung sprachkompetenter Polizist:innen unter die Erwartungen der Kollegen. Diese definieren im Einsatz sowie im sozialen Miteinander, wann und wie viel Fremdsprache angemessen ist. Zwar spare die Organisation durch die Übersetzungstätigkeiten Geld und die Kollegen Zeit (ggf. Papierarbeit), doch stützt dies weiterhin eher die Daseinsberechtigung und falle als Ressource der Sprachkompetenten nicht ins Gewicht.

Wenn davon die Rede ist, Sprachkenntnisse angemessen zu berücksichtigen, dann zielt eine solche politische Forderung nicht auf informelle Anerkennung im Kollegenkreis. Auch ist das Erkennen der Organisation, dass es einen Bedarf gibt, nicht gleichzusetzen mit dem Anerkennen der Fähigkeit. Der Mangel an institutionalisiertem Regelwerk birgt Diskriminierungsrisiken für sprachkompetente Polizist:innen. Paradoxe Weise gereicht ihnen der Vorteil letztlich zum Nachteil. Sie sind als Fremde markiert, weil sie etwas können, das die anderen nicht können. Ihre Fähigkeiten werden von der Organisation zwar als Ressource genutzt, sie »stellen aber für Personen, die diese mitbringen, in der Regel keine Ressource dar, die sie einsetzen könnten, um in der Behörde aufzusteigen« (Graevskaia 2002: 130). Vielmehr müssen sie auf reguläre, institutionelle Anerkennung verzichten, um dem Vorwurf der Bevorzugung zu entgehen. Dabei können durch innerorganisationale Regeln verursachte strukturelle Nachteile von Sprachkompetenten in der Konsequenz auch zu benachteiligender Polizeiarbeit führen, wenn z. B. nicht

übersetzt wird, obwohl es möglich wäre, weil die Beamt:innen auf die fehlende Anerkennung mit selektiver Arbeit oder Verweigerung reagieren. Dann können Diskriminierungsrisiken im Innenverhältnis der Organisation Wechselwirkungen im Außen entfalten.³⁰

5. Fazit

Es war Ziel herauszuarbeiten, welche Elemente der polizeilichen Personalauswahlverfahren typischerweise zulasten migrantischer Bewerber:innen gehen. Ich fasse zunächst die empirischen Ergebnisse zusammen und charakterisiere den Umgang der Polizeien mit Diversität in der Personalauswahl anschließend als organisierte Indifferenz. Diese führt zu strategischen Widersprüchen in der Personalauswahl.

- 1) Testelemente, die systematisch zulasten migrantischer Bewerber:innen selektieren, sind in den Phasen des Auswahlverfahrens zu finden, in denen die Bewerber:innen noch anonym gegenüber der Organisation sind. Der kognitive Leistungstest (meist am PC) selektiert überproportional zulasten migrantischer Bewerber:innen.
- 2) Die geringeren Erfolgsquoten dieser Bewerbergruppe in diesem Test gehen dabei auf jene Elemente zurück, die a) eine hohe Vertrautheit mit habituellen Erwartungen an den Polizeiberuf voraussetzen oder b) die deutsche Orthografie und Grammatik betreffen.
- 3) Mit Blick auf die Deutschkenntnisse konnte die Differenz in den Punktewerten nicht eindeutig auf geringere Bildungsvoraussetzungen zurückgeführt werden. Zwar schnitten Bewerber:innen ohne Migrationshintergrund in den polizeilichen Deutschtests in Rheinland-Pfalz im Durchschnitt besser ab. Der Unterschied in den Bildungsvoraussetzungen (gemessen an der schulischen Deutschnote und dem Schulabschluss) war jedoch zu gering, um die Leistungsdifferenz zu erklären.
- 4) In Berlin zeigte sich über Zeit (2011–2020), dass migrantische Abiturient:innen anteilig häufiger hinter den Möglichkeiten ihres Bildungsabschlusses zurückbleiben als nicht-migrantische und im mittleren Dienst beginnen, obwohl sie sich für den gehobenen Dienst beworben hatten.
- 5) In den beobachteten Auswahlgesprächen kam es teils zu stereotypisierenden Bewertungen, in denen vorurteilsgeleitete und biografisch unfaire Argumentationen zum Tragen kamen. Diese wurden über den Verweis auf ›Migrationshintergrund‹, ›Wohnort‹ oder ›Eltern bei der Polizei‹ relevant.

³⁰ Vgl. dazu das Teilprojekt und die Publikationen von Sarah Praunsmädel; <https://www.projekt-zurecht.de/arbeitbereiche/wissen-und-information/#B3>.

Die hohe Standardisierung des Verfahrens sowie der Testungen selbst unterstützt die formale Gleichbehandlung in diesem Massenverfahren, sie kann Diskriminierungsrisiken aber nicht tilgen. In den Personalauswahlgesprächen mildern standardisierte Bewertungsbögen vorurteilsgeleitete Bewertungen ab. In den computergestützten Testteilen hingegen verdeckt die Standardisierung, wenn Testelemente systematisch und überproportional zulasten migrantischer Bewerber:innen selektieren.

Die eingangs zitierte Anforderung an die Sicherheitsbehörden, nach der mittelbare Diskriminierung gegenüber dem Bewerberkreis mit Migrationshintergrund auszuschließen und sprachliche und interkulturelle Kompetenzen angemessen zu berücksichtigen sind, verlangt eine aktive Gestaltung diversitätspolitischer Ziele. Diese gestalterischen Möglichkeiten werden bezogen auf die Personalauswahl kaum genutzt, noch werden sie rechtlich als möglich eingeschätzt. Vielmehr lautet der Tenor: »Diversität ist gut, spielt aber keine Rolle.« Diversität ergibt sich, wenn die Bewerber:innen gut genug sind und es wirklich wollen; sie wird nicht hergestellt. Diese indifferente Haltung übersetzt in gewisser Weise einen rein formalen Gleichbehandlungsgrundsatz. Ohne Ansehen der Person, so die Idealvorstellung, hat jede:r die gleiche Chance die eigene Leistung unter Beweis zu stellen. Deshalb kann auch von organisierter Indifferenz gesprochen werden. Allerdings führt diese indifferente Haltung zu strategischen Widersprüche in der Rekrutierung:

- 1) Junge Menschen mit Migrationshintergrund werden von vielen Landespolizeien, der Bundespolizei und dem BKA durch Werbemaßnahmen oder im Rahmen von Netzwerkarbeit für den Polizeidienst speziell angeworben. Gleichzeitig überprüfen die Behörden nicht, ob diese Bewerbergruppe durch nicht-intendierte Verzerrungen der Tests wieder herausselektiert wird. Um einen solchen strategischen Widerspruch in der Rekrutierung zu verhindern, ist es für die Einstellungsstellen der Polizeien relevant, Diskriminierungsrisiken zu kennen, um bewusst mit ihnen umgehen zu können. Faire Testverfahren setzen die Norm des gleichen Zugangs aus Art. 33 Abs. 2 GG nicht nur im Sinne von Verfahrensgleichheit um, sondern auch im Sinne von Chancengleichheit.
- 2) Die seit den 1990er Jahren wiederholt beschriebene Erklärung, nach der Leistungsdefizite der Migrant:innen und nicht die Rekrutierungsweise der Organisation den Zugang erschweren (Holdaway 1991; Hunold 2008: 52), ist nicht nur eine negative Erwartungshaltung, die empirisch nicht so eindeutig bestätigt werden konnte. Vielmehr ergeben sich durch das *Othering* organisationskulturelle und auch institutionelle Hürden für migrantische Bewerber:innen: Die systematische Negativ-Erwartung führt dazu, dass die Berücksichtigung bspw. von zusätzlichen Sprachkenntnissen als unzulässiger Ausgleich von schlechten Leistungen gerahmt wird. So geraten positive Maßnahmen in Verdacht, den Standard zu senken. Sie werden politisch unmöglich, auch wenn sie rechtlich

möglich wären. Auch eine Orientierung an Chancengleichheit wird teilweise als Relativierung des Leistungsprinzips wahrgenommen.

- 3) Institutionelle Diskriminierung verformt und vervielfältigt sich, wenn interne Verfehlungen, die die Polizei und ihre Mitglieder betreffen (etwa rechtsextreme Chatgruppen), sich nicht in Ausschlusskriterien übersetzen. Eine indirekte Adressierung durch Verstärkung positiv besetzter Tugenden oder Kompetenzen wird dem nicht gerecht. Wenn migrantische Bewerber:innen in die Organisation rekrutiert werden sollen, um das »homogene Binnenklima« (Deutscher Bundestag 2013: 1036) zu verändern, und dann daran scheitern, dass sie bei der Überprüfung von Interkultureller Kompetenz nicht im Sinne der homogenen Deutungsmuster antworten, deutet das (nicht nur) auf einen strategischen Widerspruch in der Rekrutierung hin. Eine mangelnde Fehlerkultur führt zu Diskriminierungsrisiken für Personen, die plakativerweise als ›Lösung‹ des Organisationsproblems gewonnen werden sollten.

Literatur

- Atali-Timmer, Fatoş/Errahmouni-Rimi, Ikram (2024): Die Lösung ist das Problem: Polizei und interkulturelle Kompetenz, in: Ayça Polat/Abigail M. Joseph-Magwood (Hg.), *Alltagsrassismus – Einführung für die pädagogische Praxis und soziale Arbeit*. Stuttgart: Kohlhammer, S. 149–160.
- Behr, Rafael/Molapisi, Annelie (2022): »In der Polizei sind wir uns eigentlich alle sehr ähnlich. Eine hermeneutische Rekonstruktion der Hoffnungen, Erwartungen und Realitäten im Kontext einer Integration von Personen mit Migrationsgeschichte«, in: Antonio Vera/Rafael Behr/Martin Brussig/Anja Weiß (Hg.), *Migration und Polizei. Auswirkungen der Zuwanderung auf die Organisation und Diversität der deutschen Polizei*, S. 61–94, <https://doi.org/10.5771/9783748927273-61>.
- Deutscher Bundestag (2013): Beschlussempfehlung und Bericht des 2. Untersuchungsausschusses nach Artikel 44 des Grundgesetzes, Drucksache 17/14600.
- Deutscher Bundestag (2014): Bericht der Bundesregierung über den Umsetzungsstand der Empfehlungen des 2. Untersuchungsausschusses des Deutschen Bundestages in der 17. Wahlperiode (NSU-Untersuchungsausschuss), Drucksache 18/710, <https://dserver.bundestag.de/btd/18/007/1800710.pdf> (Zugriff: 18.09.2023).
- Ellebrecht, Sabrina (2022a): »Organisierte (In-)Differenz. Zur Bedeutung von Diversität und Repräsentation für die Polizei«, in: Daniela Hunold/Tobias Singenstein (Hg.): *Rassismus in der Polizei. Eine wissenschaftliche Bestandaufnahme*. Wiesbaden: Springer VS, S. 669–691.

- Ellebrecht, Sabrina (2022b): »Vielfalt rekrutieren: Wen rekrutiert die Polizei und wen nicht? – Öffnungsprozesse und Personalgewinnung der Polizeien in Deutschland seit den 1990er Jahren«, in: RdJB Recht der Jugend und des Bildungswesens 70/4, S. 516–536.
- Ellebrecht, Sabrina (2024): Personalauswahl für eine repräsentative Polizei, in: Tagungsband Kongress Netzwerk Demokratiestarke Polizei II, S. 169–171.
- Gomolla, Mechtilde/Radtke, Frank-Olaf (2009): Institutionelle Diskriminierung. Die Herstellung ethnischer Differenz in der Schule, 3. Auflage, Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gomolla, Mechtilde (2017): »Direkte und indirekte, institutionelle und strukturelle Diskriminierung«, in: Albert Scherr/Aladin El-Mafaalani/Gökçen Yüksel (Hg.), Handbuch Diskriminierungsforschung. Wiesbaden: Springer VS, https://doi.org/10.1007/978-3-658-10976-9_9.
- Görlich, Yvonne/Schuler, Heinz (2014): »Personalentscheidung, Nutzen und Fairness«, in: Heinz Schuler/Uwe P. Kanning (Hg.), Lehrbuch der Personalpsychologie, 3. überarb. und erweiterte Auflage, Göttingen: Hogrefe, S. 1137–1199.
- Graevskaia, Alexandra (2022): »Die Bewertung von Sprachkenntnissen im Kontext interkultureller Öffnung in der Polizei. Eine dispositivanalytische Perspektive«, in: Antonio Vera/Rafael Behr/Martin Brussig/Anja Weiß (Hg.): Migration und Polizei. Baden-Baden: Nomos, S. 95–136, <https://doi.org/10.5771/9783748927273-95>.
- Heimann, Markus (2024): Kultursensible Personalauswahl. Handreichung. https://www.bakoev.bund.de/SharedDocs/Publikationen/LG_2/Handreichung_kultursensible_Personalauswahl.pdf?__blob=publicationFile&v=2 (Zugriff 10.09.2024).
- Holdaway, Simon (1991): Recruiting a Multiracial Police Force, London: HMSO.
- Hunold, Daniela (2008): Migranten in der Polizei. Zwischen politischer Programmatik und Organisationswirklichkeit, Frankfurt a.M.: Verlag für Polizeiwissenschaft.
- Kreyenborg, Maike (2025): »Ey, ich bin jetzt nicht euer Quotentürke. Umgang der Polizei mit Diversität in der Nachwuchswerbung«, in: Sabrina Ellebrecht/Ralf Poscher/Stefan Jarolimek/Stefan Kaufmann (Hg.), Die Polizei in der offenen Gesellschaft. Zum polizeilichen Umgang mit Vielfalt und Diversität. Bielefeld: transcript, S. 67–85.
- Lang, Christine (2019): Die Produktion von Diversität in der städtischen Verwaltung. Wandel und Beharrung von Organisationen in der Migrationsgesellschaft, Wiesbaden: Springer VS.
- Leenen, Rainer/Stumpf, Siegfried/Höft, Stefan (2021): »Adverse Impact und Kulturfairness in der Personalauswahl. Empirische Befundlage und Leitlinien für die Praxis«, in: Wirtschaftspsychologie Eignungsdiagnostik für die 2020er-Jahre, 23/4, S. 94–105.

- Leenen, Wolf Rainer/Stumpf, Siegfried/Scheitza, Alexander (2014): »Interkulturelle Kompetenz in der Personalauswahl – Konzeptionalisierung und Integration in bestehende Auswahlsysteme«, in: Friederike Barié-Wimmer/Katharina von Helmolt/Bernhard Zimmermann (Hg.), *Interkulturelle Arbeitskontakte. Beiträge zur empirischen Forschung*, Stuttgart: ibidem, S. 227–258.
- Lüders, Christian (2003): »Teilnehmende Beobachtung. Stichwort«, in: Ralf Bohnsack/Winfried Marotzki/Michael Meuser (Hg.), *Hauptbegriffe qualitativer Sozialforschung*, Opladen & Farmington Hills: Barbara Budrich, S. 151–153.
- Maguer Azilis (2002): »Die Einstellung ausländischer Bewerber in den deutschen Polizeidienst«, in: *Monatsschrift für Kriminologie und Strafrechtsreform*, 85/4, S. 309–316.
- Maurer, Albrecht (1995): »AusländerInnen im Polizeidienst. Von Integrationswillen kaum eine Spur«, in Otto Diederichs (Hg.): *Hilfe, Polizei. Fremdenfeindlichkeit bei Deutschlands Ordnungshütern*, Berlin: Elefanten Press, S. 113–120.
- Mecheril, Paul (2015): »Das Anliegen der Migrationspädagogik«, in: Rudolf Leiprecht/Anja Steinbach (Hg.), *Schule in der Migrationsgesellschaft. Ein Handbuch. 1. Grundlagen – Diversität – Fachdidaktiken*, Schwalbach, Taunus: Debus Pädagogik, S. 25–54.
- Mediendienst Integration (2022): Recherche. Polizist*innen mit Migrationshintergrund, https://mediendienst-integration.de/fileadmin/Vielfalt_bei_der_Polizei_2022.pdf (Zugriff: 18.09.2023).
- Merton, Robert (1948): »The Self-Fulfilling Prophecy«, in: *The Antioch Review*, 8/2, S. 193–210.
- Moosbrugger, Helfried/Kelava, Augustin (2020): »Qualitätsanforderungen an Tests und Fragebogen«, in: Dies. (Hg.), *Testtheorie und Fragebogenkonstruktion*, 3., vollst. neu bearb., erw. u. akt. Auflage, Berlin/Heidelberg: Springer.
- Regge, Felix (2013): *Migranten in der Polizei. Eine Studie zum aktuellen Stand der Integration von Migranten in den Polizeivollzugsdienst*. Hamburg: Diplomica Verlag.
- Sacksofsky, Ute (2010): Mittelbare Diskriminierung und das Allgemeine Gleichbehandlungsgesetz, herausgegeben von der Antidiskriminierungsstelle des Bundes, https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/Expertisen/expertise_mittelbare_diskriminierung.pdf?__blob=publicationFile&v=2 (Zugriff: 18.09.2023).
- Stumpf, Siegfried/Leenen, Wolf Rainer/Scheitza, Alexander (2016): »Adverse Impact in der Personalauswahl einer deutschen Behörde: Eine Analyse ethnischer Subgruppendifferenz«, in: *German Journal of Human Resource Management*, 31/1, S. 1–28, <https://doi.org/10.1177/2397002216637289>.
- Tabbara, Tarik (2013): »Zugänge von Ausländern zur Verbeamung unter besonderer Berücksichtigung der Rechte von Drittstaatsangehörigen«, in: *Zeitschrift für Beamtenrecht* 61/4, S. 109–115.

- Weichselbaumer, Doris (2020): »Multiple Discrimination against Female Immigrants Wearing Headscarves«, in: ILR Review 73/3, S. 600–627, <https://doi.org/10.1177/0019793919875707>.
- Weichselbaumer, Doris/Schuster, Julia (2021): »The Effect of Photos and a Local-Sounding Name on Discrimination against Ethnic Minorities in Austria. A Field Experiment«, in: European Sociological Review 37/6, S. 867–882, <https://doi.org/10.1093/esr/jcab012>.
- Wisser, Laura (2025): »Drei Probleme, eine Lösung. Möglichkeiten einer repräsentativen Personalgestaltung in der deutschen Polizei«, in: Sabrina Ellebrecht/Ralf Poscher/Stefan Jarolimek/Stefan Kaufmann (Hg.), Die Polizei in der offenen Gesellschaft. Zum polizeilichen Umgang mit Vielfalt und Diversität. Bielefeld: transcript, S. 123–140.

