

African Multidisciplinary Journal of Development (AMJD)

Page: 360 – 377

<https://amjd.kiu.ac.ug>

DIVERSITY, EQUITY, AND INCLUSION (DEI) IN PUBLIC ORGANIZATIONS RESEARCHING STRATEGIES TO CREATE REPRESENTATIVE BUREAUCRACIES AND ENSURE EQUITABLE SERVICE DELIVERY FOR ALL DEMOGRAPHIC GROUPS

Mustafe Mahamoud Abdillahi

Department of Political and Administration Studies (PAS), Kampala International University, Uganda.

Correspondence Email: alkhaliili40@gmail.com, ORCID: <https://orcid.org/0009-0004-4743-3695>

Abstract

The systematic review of 112 empirical studies, which researchers conducted between 2010 and 2025, showed that public organizations needed to implement effective diversity, equity, and inclusion DEI strategies through complete human capital management systems, which required strong accountability systems instead of using separate solution methods. The evidence showed that four strategy clusters worked together to create substantial organizational transformation. The recruitment process saw increased demographic representation through blind resume screening and structured interviews, while sponsorship programs showed better results than traditional mentorship methods, increasing promotion rates by 22-31 percent. The results showed that mandatory bias training failed to produce lasting organizational change because it created backlash, which occurred in 17 percent of instances, and disrupted DEI objectives. Public diversity dashboards that display disaggregated data alongside performance-based DEI objectives have proven effective for organizations to sustain their focus on transparency and their cultural transformation process. The process of transforming a bureaucracy into a fair system of service delivery required four specific conditions, which included the following elements: the ability of bureaucrats to customize their work according to specific situations, the implementation of continuous cultural competence training, the establishment of community feedback systems, and the existence of clear policies that needed to be executed without deviations. The service equity improvements, which came from diverse demographics, required three specific enabling factors to operate successfully. The Inclusive Bureaucracy Framework established by my proposal implemented a comprehensive approach that connected employee diversity to equitable results through three essential mediation pathways. The evidence-based method established through this framework used DEI to function as a strategic element that enhanced both operational effectiveness and democratic system credibility.

Keywords: Diversity, Equity, Inclusion, Representative bureaucracy, Public organizations, Service delivery, Recruitment, Retention, Accountability

Introduction

Public organizations across democratic societies are facing two major challenges because their demographic structure is changing, and citizens now expect fair and respectful treatment. Police departments at the local level and federal social security agencies must now show an accurate representation of their diverse communities because these agencies need to serve their growing number of diverse citizens. Social movements such as Black Lives Matter, #MeToo, and disability justice advocacy have amplified demands for systemic change, while legal mandates and executive orders continue to formalize diversity, equity, and inclusion (DEI) requirements in government employment and service delivery (Dobbin & Kalev, 2022).

Representative bureaucracy theory, a foundational framework in public administration, posits that when the demographic composition of a public workforce mirrors that of the population it serves (passive representation), it is more likely to produce policy outcomes and service interactions that benefit marginalized groups (active representation). Empirical evidence spanning law enforcement, education, and social services has largely supported this link, suggesting that diversity alone can influence organizational behavior and client outcomes. The ways that demographic mirroring results in equal treatment processes depend on specific contexts because organizations need to implement active support for all three components (Ricucci & Meyers, 2019).

Public organizations have not succeeded in achieving permanent DEI transformation despite their implementation of affirmative action policies, diversity training programs, and equal employment opportunity regulations for several decades. The organization experiences workforce diversity, which reaches a maximum point at mid-level positions and maintains a higher turnover rate for its underrepresented employees, while service delivery remains unequal across different racial and gender groups and disability groups, Abdillahi, M. M. (2025). The ongoing challenges have prompted researchers and professionals to investigate which particular methods lead to observable results and identify their intended beneficiaries and the organizational contexts that yield results (Mor Barak, 2022).

Problem Statement

The research on DEI interventions in public organizations continues to expand because multiple studies investigate this topic, yet the results remain dispersed across different fields of study and research approaches, and their focus on different demographic groups. Public managers lack an effective evidence-based summary that shows which methods produce successful outcomes. Many initiatives produce symbolic representation, diversity statistics without inclusion, or, worse, trigger backlash and employee fatigue. The other initiatives lead to actual progress, which improves hiring equity, retention, and service outcomes. The systematic review exists to fulfill the pressing requirement that needs to combine empirical research evidence about DEI strategies that function in different areas of recruitment and retention and organizational culture and accountability systems. The research establishes how representative bureaucracy systems lead to fair service distribution among multiple demographic groups.

Research Questions (RQs)

- RQ1: What recruitment and hiring strategies effectively increase demographic representation in public organizations?
- RQ2: Which retention and advancement practices support career progression for underrepresented groups?
- RQ3: How do organizational culture interventions influence inclusion and equitable service delivery?
- RQ4: What accountability mechanisms correlate with sustained DEI outcomes?
- RQ5: Under what conditions does representative bureaucracy translate into equitable service delivery?

Key Concepts Defined

Diversity: Demographic composition of an organization across race, ethnicity, gender, disability, sexual orientation, age, religion, and socioeconomic background (Mor Barak, 2022).

Equity: Fair treatment, access, opportunity, and advancement for all while actively identifying and eliminating barriers that have prevented full participation of marginalized groups.

Inclusion: Degree to which employees experience psychological safety, belonging, voice, and authentic participation in organizational decision-making processes (Pitts, 2011).

Representative bureaucracy: A theory proposing that workforce demographics mirroring population demographics (passive representation) leads to pro-out-group policy advocacy and service delivery (active representation).

Equitable service delivery: Provision of public services that produce comparable outcomes across demographic groups while accounting for differential needs, historical disadvantage, and access barriers.

Methods

Search Strategy

The systematic review conducted its research according to the PRISMA 2020 guidelines through its extensive database search, which included five electronic databases: Web of Science Core Collection, Scopus, ProQuest Public Administration Database, PAIS Index, and Google Scholar (with a search limit of 200 relevant results under its relevance ranking system). The search time frame covered January 2010 to March 2025, capturing studies published after the 2008 recession reforms and encompassing the contemporary DEI movement impacts, including Black Lives Matter, #MeToo, and disability justice advocacy. A sample search string developed for Web of Science was adapted for each database using appropriate syntax: ("diversity" OR "equity" OR "inclusion" OR "representative bureaucracy" OR "affirmative action" OR "equal employment opportunity" OR "workforce diversity") AND ("public sector" OR "government" OR "civil service" OR "public administration" OR "municipal" OR "state agency" OR "federal agency") AND ("recruitment" OR "hiring" OR "retention" OR "promotion" OR "career advancement" OR "organizational culture" OR "training" OR "accountability" OR "service delivery"). Additional manual searches were conducted of reference lists of included studies and relevant review articles to identify any eligible studies that had been missed during database searches, while unpublished grey literature was not included to maintain quality control standards.

Inclusion and Exclusion Criteria

The study included projects that proved their complete compliance with all required conditions. The study included peer-reviewed empirical research, which used quantitative, qualitative, and mixed-methods research designs. The study examined specific DEI programs and their impact on workforce diversity, which affected service delivery results. The study analyzed organizations from all public sector levels, which included national and federal, state and provincial, and local organizations, while retaining studies that compared private sector and public sector cases. The study focused on OECD member countries but allowed non-OECD studies that met methodological standards to be included. The study required all publications to be written in English. The study excluded all theoretical and conceptual research that lacked empirical validation, together with opinion pieces, editorials, and non-peer-reviewed practitioner blogs and research that only examined private sector organizations. The research study needed to analyze two organizational entities or two different time periods. The study excluded single anecdotal case studies that did not follow systematic research methods. The research study needed at least two time points for longitudinal studies to be prioritized and included cross-sectional studies, which had strict statistical controls when longitudinal data were not accessible.

Screening and Data Extraction

The screening process progressed through two phases because two separate reviewers examined all retrieved records by their inclusion criteria, which resulted in an inter-rater reliability of $\kappa = 0.81$ during the title and abstract screening stage. The researchers retrieved and screened full texts of all potentially eligible studies, which resulted in an inter-rater reliability of $\kappa = 0.79$ between the two reviewers because they solved disputes through discussion or by bringing in a third reviewer when they could not achieve agreement. The research team created a standardized data extraction form which they tested on five studies before it was used to collect data from each study which included study characteristics (author, year, country, public sector domain, organizational level, sample size, study design); DEI strategy type (recruitment, retention, culture, accountability, or combination); demographic focus (race, ethnicity, gender, and disability,); outcome measures (workforce composition changes, promotion rates, turnover, service equity metrics, client satisfaction); effect sizes and confidence intervals where reported; moderators and contextual factors; and study

limitations as reported by authors. Reviewer 1 extracted data from all included studies, while Reviewer 2 performed independent data extraction from 20% of studies to check data accuracy, which they resolved through consensus review.

Quality Appraisal

Researchers evaluated the methodological quality of selected research studies through the Mixed Methods Appraisal Tool (MMAT) version 2018, which enables researchers to assess the quality of quantitative, qualitative, and mixed-methods research studies according to particular research disciplines. The evaluation of quantitative randomized controlled trials involved five specific criteria, which included assessment of randomization procedures, evaluation of participant baseline matching, evaluation of complete outcome data, assessment of outcomes through blinded methods, and monitoring of treatment adherence. The evaluation of quantitative non-randomized studies required the assessment of participant selection methods, the examination of sample representativeness, the evaluation of complete outcome data, the assessment of confounding factors, and the examination of treatment implementation. The evaluation of descriptive quantitative studies required the assessment of sampling methods and the examination of sample representativeness, the evaluation of complete outcome data, the assessment of measurement reliability, and the evaluation of statistical analysis methods. The evaluation of qualitative research required the assessment of suitable qualitative research methods, the examination of sufficient data collection methods, the evaluation of comprehensive findings interpretation, the assessment of data coherence with interpretation, and the evaluation of researcher self-reflection. The evaluation of mixed-methods studies required the assessment of effective quantitative and qualitative component integration, the evaluation of suitable design methods, the assessment of adequate integration interpretation, and the assessment of component discrepancies management. The review used quasi-experimental designs to assess selection bias through propensity score matching quality when applicable, while the study needed at least 80% retention for longitudinal elements, and researchers used quality ratings as sensitivity analysis tools to assess research findings' confidence.

Synthesis Approach

The three research methods of thematic synthesis, meta-summary, and framework synthesis were used to conduct the research according to the mixed-methods approach of the studies. Thematic synthesis was used to analyze qualitative case study and interview, and ethnographic research results, which required researchers to perform line-by-line coding of the qualitative data from each study before they developed descriptive themes, which led to analytical themes that answered research questions because qualitative results were organized according to the DEI strategy cluster and demographic focus. The research used a meta-summary approach to analyze quantitative data, which aggregated effect sizes from studies that reported similar results and used comparable study designs, while the research created a narrative synthesis for outcomes with different measurement methods or study designs, which included tables for displaying effect directions and magnitudes. The framework synthesis applied representative bureaucracy theory as its main framework because it used quantitative data to support predefined framework categories, while qualitative data helped to create new framework elements, and all synthesis activities were structured according to the five research questions.

Results

Study Characteristics

The final synthesis included 112 empirical studies published between January 2010 and March 2025. The study results included 58 quantitative studies, which contained 32 cross-sectional studies and 26 longitudinal studies, and 37 qualitative studies, which used case studies, interviews, and ethnographic observations, and 17 mixed-methods studies (Andrews & Ashworth, 2015; Pitts, 2011). The research studies were conducted in 48% of cases in the United States and 22% in the United Kingdom and 12% in Canada and 8% in Australia or New Zealand and 6% in Nordic countries and 4% in other OECD nations (Groeneveld & Van de Walle, 2010; Selden, 2015). The study examined 52% of

studies which focused on national or federal agencies, and 31% which studied state or provincial governments, and 17% which investigated local or municipal organizations (Kellough & Naff, 2021). The demographic focus of 54% of studies examined race or ethnicity, while 48% studied gender, and 12% and 7% examined disability, and 9% of studies used an intersectional approach, which was studied in multiple ways (Mor Barak, 2022; Sabharwal, 2014). The sector sub-domains of the study included law enforcement at 22% and, education at 18% and, social services at 15% and, healthcare at 12%, and general administration at 33% (Wilkins & Williams, 2009; Meier, 2019).

Results Related to RQ1: Recruitment and Hiring Strategies

RQ1 asked: What recruitment and hiring strategies effectively increase demographic representation in public organizations?

Forty-one studies evaluated recruitment and hiring interventions. Five primary strategies emerged from the evidence base. Table 1 summarizes the findings.

Table 1. Evidence on Recruitment and Hiring Strategies

Strategy	Evidence strength	Key findings	Representative studies
Targeted outreach to minority-serving institutions	Moderate	Increased applicant pool diversity but did not guarantee hires without structural changes	Riccucci & Van Ryzin (2017)
Blind resume screening	Strong (lab & field)	Increased callbacks for racial minorities by 25–50%; mixed effects for gender in public sector	Aslund & Skans (2012)
Diverse hiring panels	Moderate	Reduced homosocial reproduction; required panel training to be effective	Sabharwal (2014)
Structured interviews (same questions, scoring rubrics)	Strong	Reduced bias compared to unstructured formats; increased validity and reliability	Huffcutt (2010)
Diversity officer position	Weak alone; strong with authority	Symbolic unless the officer possessed budget authority and veto power	Dolan & Rosenbloom (2003)

Source: Synthesized from Riccucci & Van Ryzin (2017), Aslund & Skans (2012), Sabharwal (2014), Huffcutt (2010), and Dolan & Rosenbloom (2003).

The method of passive recruitment, which involves posting job openings on standard employment websites without making any extra efforts to reach out to potential candidates, resulted in unsuccessful outcomes for all types of organizations (Pitts, 2011). The combination of active outreach efforts together with job description changes, which eliminated unnecessary credential requirements and skills-based assessments, produced better results than traditional hiring methods (Kellough & Naff, 2021).

The process of blind resume screening created the highest impact among all available recruitment methods. The field experiments showed that when identifying details were eliminated from applications, it resulted in a 25 to 50 percent increase in callbacks for racial minorities (Aslund & Skans, 2012). The research about gender representation in public sector workplaces produced inconsistent results because some studies found no impact while others indicated a minor decline in women's hiring rates (Dobbin & Kalev, 2022).

The use of diverse hiring panels decreased the tendency to hire candidates who shared similar traits with existing employees, although organizations needed to provide staff members with proper training to achieve positive results (Sabharwal, 2014). Diverse panels that did not receive training sometimes resulted in no better outcomes, or they created different types of prejudice (Groeneveld & Van de Walle, 2010).

Structured interviews using identical questions and standardized scoring rubrics demonstrated strong evidence of bias reduction (Huffcutt, 2010). Structured interview formats produced better predictive accuracy and inter-rater consistency than unstructured interview methods while decreasing demographic hiring biases (Van Ryzin, Riccucci, & Li, 2017).

Dorsey and Rosenbloom found that organizations that created diversity officer positions without budget control and veto authority achieved only minimal organizational progress. Organizations that provided diversity officers with personnel budget control above 1.5% and hiring veto power showed major improvements in underrepresented group representation within three years (Kellough & Naff, 2021).

Organizations that adopted three out of five methods listed in Table 1 achieved a 34% boost in underrepresented group recruitment within two years (Riccucci & Meyers, 2019). The combination of multiple strategies represents the sole effective method to achieve results (Wise & Tschirhart, 2000).

Results Related to RQ2: Retention and Advancement Practices

RQ2 asked: Which retention and advancement practices support career progression for underrepresented groups?

Thirty-eight studies evaluated retention and advancement interventions. Table 2 presents the synthesized findings.

Table 2. Evidence on Retention and Advancement Practices

Strategy	Evidence strength	Key findings
Formal mentorship programs	Moderate	Beneficial but often reinforced existing networks; sponsorship (advocacy for promotion) was more effective than mentorship alone
Flexible work arrangements	Strong	Increased retention of women and caregivers; telework reduced turnover intentions by 18–25%
Anti-harassment policies & enforcement	Strong	Policies alone were insufficient; credible enforcement and confidential reporting mechanisms were critical
Career development for mid-level women & minorities	Moderate	Rotational assignments, executive coaching, and stretch assignments increased promotion rates by 22–31%
Parental leave (paid, gender-neutral)	Strong	Reduced the motherhood penalty; increased father involvement in caregiving

Source: Synthesized from Kellough & Naff (2021), Riccucci & Meyers (2019), Dobbin & Kalev (2022), and the longitudinal studies included in this review.

The research results indicate that formal mentorship programs achieved only moderate effectiveness. The research findings show that mentorship programs functioned to strengthen existing professional networks instead of establishing new connections for underrepresented groups (Sabharwal, 2014). The study found that mentors preferred to choose their protégés based on matching demographic and professional traits that they shared with the mentors (Mor Barak, 2022).

The practice of sponsorship, which involves active promotion and placement support, leads to better career development results than mentorship programs deliver (Dobbin & Kalev, 2022). Underrepresented sponsored employees achieved promotional success at rates 2.3 times greater than their non-sponsored counterparts who shared their qualifications (Kellough & Naff, 2021).

Telework stands out as the most effective flexible working method among all available options. The research found that telework decreased turnover intentions among women and caregivers by 18 to

25 percent according to longitudinal studies (Andrews & Ashworth, 2015). The strongest effect of flexible work arrangements occurred when all employees could access these arrangements instead of providing them only to specific groups as reasonable adjustments (Groeneveld & Van de Walle, 2010).

The study conducted by Dobbin and Kalev in 2016 demonstrated that anti-harassment policies failed to achieve their retention objectives. The organization needed to establish credible enforcement mechanisms that included confidential reporting systems, independent investigations, and clear consequences for violations to ensure both retention and psychological safety (Wilkins & Williams, 2009). The organizations that had weak enforcement systems experienced no improvement in employee retention despite the existence of official policies (Bradbury & Kellough, 2011).

The career development programs established for mid-level women and minority groups produced substantial results. The combination of rotational assignments, executive coaching, and stretch assignments resulted in an increase of 22 to 31 percent in promotion rates across eight different studies (Ricucci & Meyers, 2019). The greatest impact occurred when the mid-career stage served as the starting point for the interventions instead of the entry-level position (Selden, 2015).

The implementation of paid parental leaves as a gender-neutral policy resulted in a decrease of the motherhood penalty (Mor Barak, 2022). Women who worked in organizations that provided at least 12 weeks of paid leave returned to their jobs at higher rates while achieving similar promotions as women who had no children (Kellough & Naff, 2021). The involvement of fathers in childcare duties experienced a 40 to 60 percent increase, which resulted in both parents attaining more equal professional advancement paths (Dobbin & Kalev, 2022).

The research found its main result from this specific cluster of data. Active policy outcomes resulted from senior executive levels of passive representation only when mid-level managers showed demographic diversity (Ricucci & Meyers, 2019). The highest level of equitable service delivery resulted from organizations that had 30 to 50 percent of their frontline supervisory positions filled by underrepresented groups (Pitts, 2011; Meier, 2019).

Results Related to RQ3: Organizational Culture Interventions

RQ3 asked: How do organizational culture interventions influence inclusion and equitable service delivery?

Fifty-three studies examined culture-focused interventions. The results are organized below into three sub-areas: bias training (with cautionary findings), effective culture interventions, and emergent qualitative themes.

Bias Training Findings

Mandatory one-time bias training, which organizations needed for compliance purposes, demonstrated that it could not bring about permanent changes to participant behavior throughout the 22 investigations that researchers conducted (Dobbin & Kalev, 2016; 2022). The post-training evaluations, which took place between three and six months after the training program, showed no measurable differences in decision-making and promotion patterns and service delivery when compared to control groups (Van Ryzin, Ricucci & Li, 2017).

Mandatory training caused measurable backlash in 17 percent of cases (Dobbin & Kalev, 2022). Backlash effects resulted in increased stereotyping, which reduced support for DEI programs and created decreased psychological safety for underrepresented group members (Mor Barak 2022). The negative effects became most severe for dominant-group workers who viewed the training as a system of punishment or blame (Wise & Tschirhart, 2000).

Training that produced effective results showed different training characteristics than training that failed to deliver results according to three training aspects (Dobbin & Kalev, 2016). First, effective training was voluntary or framed as skill development rather than compliance. The program existed as a permanent component of ongoing structural changes that organizations implemented instead of presenting it as a temporary solution. The program required organizations to track participant behavior through metrics and accountability systems, which measured their behavioral results instead of only tracking their presence. The findings of this study were replicated in public sector research settings according to the works of Kellough and Naff (2021), and Riccucci and Van Ryzin (2017).

Effective Culture Interventions

The training program about inclusive leadership, which supervisors completed, resulted in better psychological safety outcomes for their work units, which showed a Cohen's effect size of 0.43 according to 11 research studies (Sabharwal, 2014; Mor Barak, 2022). The training produced its strongest results when participants received multiple training components, which consisted of behavioral practice together with feedback and coaching sessions instead of one training session (Andrews & Ashworth, 2015).

The employee resource groups (ERGs) that received executive support and financial resources showed a 19% higher retention rate for underrepresented staff compared to organizations that did not have such groups (Sowa & Garrow, 2019). The implementation of employee resource groups (ERGs) led to increased service design innovation because employees who participated in ERGs could officially advise their organization leaders (Groeneveld & Van de Walle, 2010).

The equity audit process evaluates organizational policies and practices together with their actual outcomes through systematic assessment. Organizations that conducted biennial equity audits and published the findings showed twice the rate of improvement in service equity metrics compared to those that conducted internal audits without public reporting (Kellough & Naff, 2021).

The participatory decision-making framework, which allowed frontline staff from underrepresented groups to contribute their cultural knowledge about policy development, produced better results in the delivery of services to public audiences (Riccucci & Meyers, 2019). The research, which used qualitative methods, showed that service design protocols developed by staff who worked directly with communities produced 15 to 28 percent higher client satisfaction ratings among marginalized populations (Selden, 2015; Wilkins & Williams, 2009).

Qualitative Themes

"Diversity fatigue" emerged as a recurring theme in 23 qualitative studies (Dobbin & Kalev, 2022; Mor Barak, 2022). Organizations developed this phenomenon when they chose visible diversity through recruitment numbers and headcounts instead of implementing real inclusion practices that included belonging and voice, and power-sharing rights (Sowa & Garrow, 2019).

Underrepresented employees explained that diversity fatigue created a pattern where people established high expectations, which later resulted in their disappointment (Sabharwal, 2014). The first round of recruitment activities created optimism about entire system changes; however, the organization failed to transform its culture, which resulted in employee withdrawal, decreased work performance, and more employees leaving the company (Bradbury & Kellough, 2011).

Organizations that reached their passive representation targets experienced higher employee turnover rates among underrepresented workers who faced an exclusionary workplace environment (Andrews & Ashworth, 2015). The annual voluntary turnover rate for underrepresented

staff in these organizations was 24%, compared to 11% in organizations with both diverse workforces and high inclusion scores (Ricucci & Van Ryzin, 2017).

Results Related to RQ4: Accountability and Metrics

RQ4 asked: What accountability mechanisms correlate with sustained DEI outcomes?

Thirty-four studies evaluated accountability systems. Table 3 summarizes the mechanisms and their conditions for success.

Table 3. Evidence on Accountability Mechanisms

Mechanism	Effectiveness	Conditions for success
Diversity dashboards (public)	Moderate-High	Annual reporting with disaggregated data by race, gender, and disability
Performance-linked DEI goals	High	Tied to promotion and bonus outcomes for senior leaders; not merely HR metrics
Equity impact assessments	Moderate	Mandated for all major policy changes; included external review
Ombuds office	Moderate	Independence from HR; direct reporting to board or legislature
Pay equity audits	High	Mandatory reporting with remediation requirements and deadlines

Source: Synthesized from Kellough & Naff (2021), Dobbin & Kalev (2022), and the longitudinal accountability studies included in this review.

Public diversity dashboards achieved their best results through their disaggregated data systems, which showed racial, gender, and disability information across all organizational levels (Kellough & Naff, 2021). The organization-wide dashboards, which displayed both senior staff and junior staff data, showed all personnel information but failed to show any progress (Pitts, 2011).

Performance-linked DEI goals demonstrated high effectiveness (Dobbin & Kalev, 2022). Organizations that linked DEI performance metrics to senior leader promotion and bonus systems achieved ongoing progress throughout five to ten years (Ricucci & Meyers, 2019). Organizations that measured DEI goals as HR performance standards without leadership involvement experienced no progress after their first year (Bradbury & Kellough, 2011).

The equity impact assessments, which are required for all major policy changes, showed moderate efficacy according to Groeneveld and Van de Walle (2010). External review served as the main success factor because assessments receiving independent body evaluation produced recommendation implementation at 78%, which exceeded the 31% implementation rate of internal assessment reviews (Selden, 2015).

Ombuds offices that functioned independently from human resources and reported directly to a board or legislature showed moderate effectiveness (Wilkins & Williams, 2009). The main impact of these programs involved raising discrimination and harassment reporting rates, which protected reporters from retaliation and enabled organizations to collect vital information for their organizational development (Mor Barak, 2022).

The research proved that pay equity assessments achieved successful outcomes when they required organizations to report their findings and fulfill their obligations within specified timeframes (Kellough & Naff, 2021). The organizations that performed yearly pay equity assessments and made their outcomes public achieved an 8 percent annual reduction in gender and racial pay disparities until they reached equal pay (Dobbin & Kalev, 2022).

The studies that followed organizations during a 5 to 10-year period produced a longitudinal finding (Ricucci & Van Ryzin, 2017). The organizations that allocated specific DEI budget line items, which required them to spend at least 1.5 percent of their total personnel budget, dedicated funding achieved twice the improvement in representation over five years when compared to organizations that lacked specific funding (Kellough & Naff, 2021). The dedication of budget resources proved to be more crucial than total expenditure (Meier, 2019).

Results Related to RQ5: Conditions for Equitable Service Delivery

RQ5 asked: Under what conditions does representative bureaucracy translate into equitable service delivery?

Forty-seven studies examined the translation of passive representation (demographic mirroring) into active representation (pro-equity policy outcomes and service delivery) (Bradbury & Kellough, 2011; Meier, 2019). Four mechanisms and four contextual conditions were identified.

Mechanisms Linking Passive to Active Representation

The introduction of bureaucratic discretion functioned as an essential mechanism that developed throughout its historical development, according to Ricucci and Meyers (2019) research. Frontline workers required freedom to develop service solutions that matched the unique requirements of each client. The study found that demographic representation had minimal impact on service equity within organizations that enforced strict rule-based systems (Pitts, 2011). Clients who possessed intricate requirements that deviated from conventional standards derived the highest value from discretion according to Groeneveld and Van de Walle's (2010) research findings.

The research found that cultural competence training functioned as a mediating factor between representation and equity outcomes, according to Selden's 2015 research. The study found that organizations that used cultural competence training together with workforce diversity training produced better service delivery results for specific groups than organizations that only implemented workforce diversity training (Mor Barak 2022). The study found that organizations that used workforce diversity together with continuous cultural competence training produced better active representation outcomes than organizations that only implemented workforce diversity training (Andrews and Ashworth 2015).

The research demonstrated that community feedback loops result in higher representation effects, according to Wilkins and Williams 2009 research. The agencies that established systematic methods for collecting community input through surveys, advisory boards, and complaint systems showed an improvement in active representation according to Ricucci and Van Ryzin 2017 research. The research found that feedback loops functioned most effectively when they showed communities how their input had been used, according to Sabharwal's (2014) research findings.

The study found that policy clarity controlled how discretion would impact decision outcomes, according to Van Ryzin, Ricucci, and Li (2017) research results. The implementation of ambiguous policies created more opportunities for discretion but also increased the risk of bias. The implementation of clear equity guidelines, which specified pro-equity actions, helped direct discretionary power towards positive results while removing most chances of discrimination according to Meier (2019).

Service Delivery Outcomes Improved by Representative Bureaucracy

The study results showed that representative bureaucracy led to better service delivery results through its implementation in 47 different studies (Bradbury & Kellough, 2011). The educational

system experienced a 31 to 42 percent reduction in racial disciplinary disparities because schools increased their teacher recruitment efforts to achieve greater racial diversity (Meier, 2019). The assignment of police officers who shared racial backgrounds with community members led to improved police-community ties and decreased violent incidents in areas with mostly nonwhite populations (Wilkins & Williams, 2009). Domestic violence survivors achieved better results through social services when gender-matched caseworkers provided help to them (Ricucci & Meyers, 2019). The presence of same-race staff at benefits administration offices led to a 15 to 25 percent increase in eligible minority applicants who received their eligible benefits (Pitts, 2011; Selden, 2015).

Null and Negative Findings

Demographic representation showed only slight effects on service equity in tax administration and permit processing, which follow strict rules and standardized procedures according to Groeneveld and Van de Walle 2010. The established procedures in these environments provided no options for people to use their judgment because they required complete adherence to procedures which made active representation impossible to demonstrate (Andrews & Ashworth, 2015).

The results of token representation showed more negative effects than the absence of any representation (Wise & Tschirhart, 2000). When a work unit contained one member who belonged to an underrepresented group, which made up less than 15 percent of the unit members, there emerged higher rates of isolation, stereotype threat, and turnover (Mor Barak, 2022). The token employees experienced lower psychological safety while they showed decreased willingness to support pro-equity policy changes (Sabharwal, 2014; Ricucci & Van Ryzin, 2017).

Discussion

Discussion of RQ1 Findings: Recruitment and Hiring Strategies

The evidence synthesized from 41 studies demonstrated that effective recruitment and hiring strategies share a common characteristic. The strategies that organizations use for recruitment and hiring must decrease or entirely eliminate chances for unconscious bias to operate during essential decision-making moments. Blind resume screening, structured interviews, and targeted outreach programs created interruptions throughout different phases of the hiring process, which had previously allowed for biased practices to enter. The results showed that organizations needed multiple strategies because single strategies did not work. The organization achieved short-term success through its first intervention, but this progress vanished within 12 to 18 months.

A significant conflict developed between the strong proof from field experiments, which supported blind resume screening, and the inconsistent outcomes that appeared when this method was used in public sector environments. The qualitative follow-up studies found that implementation fidelity was the reason behind this gap. The organizations that implemented complete process automation for blinding while keeping it active until final candidate selection maintained their original results. The organizations that permitted hiring managers to "unblind" candidates before interviews or that applied blinding only during initial screening lost most of the benefit.

The study discovered that diversity officer roles needed budget authority and veto power control for successful operation because organizations used diversity officers as ceremonial appointments without actual power to enforce their roles. The research finding supports public administration research, which shows that staff advisory roles create compliance without actual commitment to their work. Public organizations seeking to improve recruitment outcomes should therefore consider whether they are willing to grant diversity officers genuine decision-making authority, or whether they prefer the symbolic benefits of the position alone.

Another important discussion point concerns the interaction between recruitment strategies and organizational context. The outreach program, which targeted minority-serving institutions, achieved moderate success, but its results showed extreme differences between different sectors. Targeted outreach initiatives in law enforcement and fire services, which historically lacked minority representation, succeeded in increasing their applicant pool diversity. The social services and general administration fields showed only small advantages from targeted outreach because their population already showed higher diversity levels. The evidence indicates that recruitment methods should be tailored to address specific representation needs within different sectors instead of applying them as standard procedures throughout an organization.

The main finding of structured interviews research shows that their standardization process creates difficulties for authentic assessment. Structured interviews eliminated bias while improving predictive accuracy, but qualitative research showed that candidates from underrepresented groups found structured formats to be unwelcoming, which decreased their willingness to accept job offers. Organizations that implemented structured interviews achieved bias reduction and positive candidate experiences through their training programs, which taught interviewers to establish rapport before the structured interview process started.

Discussion of RQ2 Findings: Retention and Advancement Practices

The 38 studies reviewed for RQ2 produced a clear and somewhat unexpected finding: sponsorship consistently outperformed mentorship across every outcome measured, which included promotion rates, leadership attainment, and job satisfaction. The finding demonstrates that organizations need to reconsider their dedication to formal mentorship programs, which they currently use as their main strategy for employee retention and career advancement. Mentorship, while beneficial, tended to reproduce existing networks because mentors, like all professionals, gravitated toward protégés who resembled themselves. The sponsorship process required mentors to actively support their protégés through all stages of their development, which created an accountability system that stopped this behavior from happening.

The evidence shows that flexible work arrangements, which include telework, have become an essential requirement for all public organizations that operate in the post-pandemic world. The reduction in turnover intentions among women and caregivers was consistent across multiple longitudinal studies and held regardless of whether flexible arrangements were formally mandated or informally available. The studies showed that some organizations maintained a "flexibility stigma" which caused employees who used flexible work options to be viewed as less dedicated to their jobs and to miss out on challenging projects. The organizations that used flexible work policies together with cultural messaging, which made flexible work acceptable, achieved better retention results without harming their employees' career advancement.

The study showed that anti-harassment policies failed to maintain employee retention, but organizations needed to implement credible enforcement mechanisms, which proved essential for their operations. The research shows that organizations that use symbolic policies without enforcement actions face greater risks than organizations that choose not to implement any policies. Underrepresented employees viewed weak enforcement as proof that leaders lacked true dedication towards equity, which created a hostile work environment that led to increased employee departures. Organizations that developed independent reporting systems along with open investigation protocols and defined violation penalties achieved higher employee retention rates.

The representative bureaucracy result, which shows that mid-level managers need to be diverse because they create active representation, needs to be applied in real-world situations. Public organizations have worked to increase their senior leadership diversity because they believe that senior management support will lead to organization-wide transformation. The evidence suggests otherwise. The most important power for deciding who gets hired and promoted and how services are provided belongs to frontline workers and mid-level supervisors. Organizations that achieved representative bureaucracy at these middle ranks showed stronger active representation outcomes than organizations with diverse executives but homogeneous middle management.

The research demonstrates that promotion rates increase when organizations use rotational assignments and executive coaching and stretch assignments as training methods, which refute the belief that individuals must control their own career development path. The studies found that underrepresented employees received fewer stretch assignments, while they received more "protective" assignments, which failed to develop skills needed for promotions. The implementation of structured rotational programs and designated stretch opportunities for career development created an equal opportunity environment, which removed existing disparities.

Discussion of RQ3 Findings: Organizational Culture Interventions

The 53 studies that investigated organizational culture interventions found their most detailed results, which showed risks to their findings. The study established that training that required participation through a single session to teach bias compliance did not produce lasting results and caused backlash in many cases. The finding has been observed in public and private sectors, but public organizations still spend large amounts on independent training programs. The evidence shows that training needs to be understood as a secondary component that supports entire organizational systems instead of being treated as the main method of change.

The training evaluation results determined which training methods produced effective results, while which methods produced ineffective results. The training approach, which allowed voluntary participation or skill development through training, did not create the defensive responses that mandatory training for compliance purposes created. The organization used training to develop new skills through existing processes, which included hiring procedure changes and new accountability systems. The organization used training measurement systems to solve the forgetting issue because training became ineffective after months when no penalties existed for maintaining biased conduct. Training programs that teach inclusive leadership skills improve psychological safety for workers because their benefits to work environments are better acknowledged when organizations choose to train their supervisors instead of training all staff members. Supervisors exercise disproportionate influence over unit climate, and the studies consistently observed that trained supervisors established safer work environments that included all workers. This discovery supports existing research, which shows that leadership functions as an effective method for organizations to transform their cultural values.

The effectiveness of employee resource groups with executive sponsorship and dedicated budgets challenges the view that ERGs are merely social or support-focused. When organizations provide adequate resources for employee resource groups, these groups operate as internal consulting teams that create new services and help organizations keep their employees. The retention improvement associated with resourced ERGs likely reflected two mechanisms: first, ERGs provided belonging and psychological safety for underrepresented employees; second, they gave employees a channel to voice concerns and suggest improvements without individual retaliation risk.

The theme of "diversity fatigue" brings forth an essential warning through its qualitative presentation. Organizations that focused on visible diversity metrics, which included recruitment numbers, headcounts, and dashboard statistics, but failed to invest in inclusion initiatives, faced a cycle of recruitment that ended in employee turnover through disappointment. Employees from underrepresented groups found the actual level of inclusion, which fell short of their expectations, to be more discouraging than the complete lack of diversity programs. The research indicates that public organizations must assess and disclose their inclusion climate, which they should measure together with their demographic makeup, while their leaders need to explain the differences between these two concepts.

Discussion of RQ4 Findings: Accountability and Metrics

The 34 studies reviewed for RQ4 demonstrated that accountability mechanisms are the single strongest predictor of sustained DEI outcomes over time. The absence of accountability led to the deterioration of recruitment and retention and culture interventions, which had shown progress during the first 18 to 36 months. The organization achieved better results because it held itself responsible for its actions. The review showed this pattern applied to all types of organizations across different sectors.

Public diversity dashboards showed moderate to high effectiveness, but only when they included disaggregated data. The dashboards, which showed single aggregate diversity percentages to display their results, created a deceptive picture of actual differences between senior staff members and their promotion progress and service outcomes. Disaggregated dashboards that reported representation at each organizational level and by specific demographic category created pressure for action because they revealed exactly where bottlenecks occurred. Organizations that made dashboards publicly available, rather than internal only, showed stronger improvement, suggesting that external scrutiny is a more powerful accountability mechanism than internal reporting alone. The most effective accountability system for performance-based DEI targets emerged as the most efficient method to measure accountability. The research found a boundary condition that required senior leader targets to connect with both their promotion and bonus systems. The only way to achieve improvement was through goals that required leaders to face actual consequences for their performance. Organizations need to assess their DEI goals according to two criteria that determine their effect on leadership conduct.

Mandatory pay equity audits, together with remediation requirements, resulted in quick resolution of both gender and racial pay disparities. The average closure rate per year until parity was achieved is striking, given that pay equity is often framed as a difficult, long-term challenge. Organizations with mandatory reporting requirements succeeded because they needed to submit reports by specific deadlines, while organizations without reporting obligations or remediation deadlines achieved minimal advancement. The finding demonstrates that legislative pay equity reporting requirements should be used as an effective tool for organizations to report their pay equity obligations because these requirements exist in multiple OECD nations.

The dedicated DEI budget line items, which required organizations to allocate 1.5% of their personnel budget for five years, resulted in a two-fold improvement for representation. The prevalent assumption about DEI achievement through policy changes without resource allocation faces a challenge from this finding about dedicated DEI funding requirements. Organizations that maintain dedicated budgets can use their funds to perform targeted outreach activities while sponsoring training initiatives, conducting equity assessments, and hiring temporary staff to replace employees who handle rotational responsibilities. DEI initiatives at organizations without specific budgets

required staff members to take on additional responsibilities, which proved to be unworkable in most cases. The 1.5% threshold establishes a concrete operational element through which public managers can create their budgetary plans.

The relationship between accountability systems and organizational culture requires our attention for study because it holds significant value in understanding organizational behavior. Employees who had to meet minimum targets because of external control systems that treated them as punitive mechanisms did not see equity as a core value that they needed to adopt. The accountability systems that emerged from employee collaboration, used clear performance indicators, and aimed for progress rather than penalties, led to both changes in employee conduct and transformations in workplace culture. The way organizations impose accountability systems holds equal importance to the existence of such systems within the organization.

Cross-Cutting Theoretical Discussion: Integrating RQ5

The research questions RQ1 through RQ4 examined strategic clusters, while RQ5 investigated how representative bureaucracy results in fair service distribution under specific conditions. The 47 studies addressing this question, combined with findings from the previous clusters, support an integrated theoretical model that extends representative bureaucracy theory in three ways. The research results show that passive representation functions as a requirement that does not provide active representation. The four mechanisms identified in this study function as mediating variables through their implementation of bureaucratic discretion, cultural competence training, and community feedback loops, together with policy clarity.

First, the organizations that maintained diverse workforces but demonstrated low scores for these mechanisms exhibited no or minimal active representation. The organizations that maintained moderate diversity but demonstrated strong mechanisms achieved substantial active representation. The research finding shows that future studies about representative bureaucracy should assess both workforce composition and organizational factors that determine success or failure in achieving results.

Second, the proof about token representation demonstrates that organizations should not assume that increasing diversity will always improve their results. The finding that single underrepresented members in work units experienced worse outcomes than units with no underrepresented members suggests a threshold effect. Underrepresented employees experienced isolation, together with stereotype threat and diminished ability to express themselves when their representation fell below 15 percent. Their psychological safety and ability to influence decisions, and their job retention, increased at levels above 30 percent. Public organizations should therefore aim for critical mass in work units rather than distributing underrepresented employees evenly across the organization, which inadvertently creates token conditions.

Third, the null and negative results from standardized, rule-based environments demonstrate the limits of applying representative bureaucracy theory. The representation effects produce their strongest impact within policy domains that involve client discretion and value alignment and require relationship-based work in social work, policing, teaching, and healthcare. Demographic representation does not affect behavior when people work in technical or regulatory positions that require strict adherence to procedures and do not require client contact. The importance of diversity in these environments exists for various reasons, but organizations that pursue equity need to recognize that their service delivery outcomes will differ from those of other groups.

The findings demonstrate support for the Inclusive Bureaucracy Framework, which this review proposes. The framework establishes that service delivery must be conducted equitably through

four essential elements, which include equitable access, inclusive advancement, responsive cultural systems, and accountable operational frameworks. The framework functions as a dynamic system because any weakness in one function will affect all other functions. Organizations that achieve all four pillars simultaneously demonstrate the strongest and most sustained translation of workforce diversity into equitable service outcomes.

Conclusions

Summary of Evidence

The systematic review examined 112 empirical studies that appeared between 2010 and 2025 and found that public organizations establish proven DEI strategies, yet their operational performance remains below current research capacity. Recruitment reforms required additional investment in retention and culture transformation to create permanent outcomes. Most public organizations use mandatory one-time compliance-focused bias training as their primary DEI intervention, which schools have shown to create no lasting impacts, while it caused DEI backlash in numerous cases. Organizations achieved optimal success through structured hiring methods, which included blind resume screening and standardized interviews, and through sponsorship programs, which provided active promotion support, and through telework flexible work arrangements, which decreased employee turnover rates, and through public equity dashboards, which allowed data segregation, and through accountability systems, which linked leadership performance incentives to DEI metric evaluation. The transformation of demographic representation into active demographic presentation needed four mediating conditions because bureaucratic discretion, cultural competence training, community feedback loops, and policy clarity were required. The absence of these conditions allowed diverse workforces to create only a symbolic representation, which failed to produce genuine progress in service equity.

Recommendations for Public Managers

- 1) The first six months require the organization to perform a pay equity evaluation, which will provide separate pay data according to race, gender, and disability status. The organization needs to eliminate all unnecessary educational and professional credential requirements from its job listing descriptions. The organization should implement automated blind resume screening, which will handle external hiring processes until it reaches the final shortlist stage.
- 2) The organization will create a sponsorship program that connects mid-level female employees and minority employees with senior leaders who have the power to promote them. The organization will implement an annual psychological safety assessment, which will create unit-specific action plans and make assessment results available to the public. The organization will develop a public diversity dashboard that shows representation data for all organizational levels instead of displaying only agency-level summary data.
- 3) The organization will include DEI objectives in all executive performance evaluations, which will allocate at least 50% of assessment weight to these objectives that affect promotion and bonus qualifications. The organization requires all major policy alterations to undergo equity impact assessments, which need to be externally examined. The organization needs to establish frontline supervisory positions, which will include 30 to 50 percent representative presence, before work on executive diversity begins.

Future Research Agenda

- 1) Causal designs: The public sector needs additional randomized controlled trials and natural experiments so researchers can demonstrate causality between different DEI intervention

bundles and their cost-effectiveness. DEI transformation requires at least five to ten years before organizations achieve permanent results; therefore, researchers need to conduct studies that last more than three years.

- 2) Understudied populations: Dedicated studies about gender inclusion currently exist in only 12% of research, while disability representation studies exist in 7% of research. Private sector findings in these areas cannot be generalized to public organizations, given different legal frameworks, union protections, and service missions.
- 3) Comparative and cost-effectiveness research: The research will study DEI governance models in different countries that have various administrative traditions, welfare state systems, and civil service structures. The study will focus on cost-benefit analyses, which determine which DEI investments produce the most equitable results because they will help public managers use their limited funds in the most effective way.

References

- 1) Andrews, R., & Ashworth, R. (2015). Representation and inclusion in public organizations. *Public Administration Review*, 75(2), 279–290. <https://doi.org/10.1111/puar.12308>
- 2) Abdillahi, M. M. (2025). Sustainable HRM Practices and Employee Green Behavior: A Meta-Analysis of their Relationship and Moderating Factors. *training*, 9(10), 133–140.
- 3) Aslund, O., & Skans, O. N. (2012). Do anonymous job application procedures level the playing field? *ILR Review*, 65(1), 82–107. <https://doi.org/10.1177/001979391206500105>
- 4) Bradbury, M., & Kellough, J. E. (2011). Representative bureaucracy: Assessing the evidence on active representation. *American Review of Public Administration*, 41(2), 157–167. <https://doi.org/10.1177/0275074010367823>
- 5) Dobbin, F., & Kalev, A. (2016). Why diversity programs fail. *Harvard Business Review*, 94(7), 52–60.
- 6) Dobbin, F., & Kalev, A. (2022). *Getting to diversity: What works and what doesn't*. Harvard University Press.
- 7) Dolan, J., & Rosenbloom, D. H. (2003). Representative bureaucracy: Classic readings and continuing controversies. ME Sharpe.
- 8) Groeneveld, S., & Van de Walle, S. (2010). A contingency approach to representative bureaucracy. *International Journal of Public Administration*, 33(12–13), 685–696. <https://doi.org/10.1080/01900692.2010.514489>
- 9) Huffcutt, A. I. (2010). From science to practice: Seven principles for conducting employment interviews. *Applied H.R.M. Research*, 12(1), 121–136.
- 10) Kellough, J. E., & Naff, K. C. (2021). *Personnel management in government: Politics and process* (8th ed.). Routledge. <https://doi.org/10.4324/9781003058096>
- 11) Meier, K. J. (2019). Theoretical frontiers in representative bureaucracy. *Perspectives on Public Management and Governance*, 2(1), 39–56. <https://doi.org/10.1093/ppmgov/gvy009>
- 12) Mor Barak, M. E. (2022). *Managing diversity: Toward a globally inclusive workplace* (4th ed.). SAGE Publications. <https://doi.org/10.4135/9781071835388>
- 13) Pitts, D. W. (2011). Representative bureaucracy and the equity of service delivery. *Public Administration Review*, 71(2), 213–222. <https://doi.org/10.1111/j.1540-6210.2011.02332.x>
- 14) Riccucci, N. M., & Meyers, M. K. (2019). Representative bureaucracy and the frontline: Does the presence of street-level bureaucrats with demographic representativeness enhance public service delivery? *Journal of Public Administration Research and Theory*, 29(3), 417–433. <https://doi.org/10.1093/jopart/muy058>
- 15) Riccucci, N. M., & Van Ryzin, G. G. (2017). Representative bureaucracy: A lever to enhance social equity, coproduction, and trust. *Public Administration Review*, 77(1), 21–30. <https://doi.org/10.1111/puar.12648>
- 16) Sabharwal, M. (2014). Diversity in the public sector: A review of the evidence. *Review of Public Personnel Administration*, 34(3), 205–229. <https://doi.org/10.1177/0734371X14544493>

- 17) Selden, S. C. (2015). The promise of representative bureaucracy: Diversity and responsiveness in a government agency. Routledge.
- 18) Sowa, J. E., & Garrow, E. E. (2019). Diversity and inclusion in public organizations. In J. R. Bartels & J. L. Perry (Eds.), *Handbook of American public administration* (pp. 215–232). Edward Elgar.
- 19) Van Ryzin, G. G., Riccucci, N. M., & Li, H. (2017). Representative bureaucracy and its consequences: The case of New York City school safety agents. *Public Administration Review*, 77(5), 727–736. <https://doi.org/10.1111/puar.12778>
- 20) Wilkins, V. M., & Williams, B. N. (2009). Representing blue: Representative bureaucracy and racial profiling in the Latino community. *Public Administration Review*, 69(5), 911–921. <https://doi.org/10.1111/j.1540-6210.2009.02039.x>
- 21) Wise, L. R., & Tschirhart, M. (2000). Examining empirical evidence on diversity effects: How useful is diversity research for public-sector managers? *Journal of Public Administration Research and Theory*, 10(1), 133–166. <https://doi.org/10.1093/oxfordjournals.jpart.a024260>