



Article

Human Resource Capacity and the Self-Assessment Gap in Devolved Public Social Welfare Offices: Evidence from Local Social Welfare and Development Officers in the Philippines

Richard A. Dizon¹, Joseph C. Barrera^{2,*}

¹Graduate School, Lyceum Northwestern University, Dagupan City, Pangasinan 2400, Philippines

²BS Social Work Department, Pangasinan State University –Lingayen Campus, Lingayen, Pangasinan 2401, Philippines

*Corresponding author: Joseph C. Barrera, jbarrera.lingayen@psu.edu.ph

CITATION

Dizon RA, Barrera JC. Human Resource Capacity and the Self-Assessment Gap in Devolved Public Social Welfare Offices: Evidence from Local Social Welfare and Development Officers in the Philippines. *Metaverse*. 2026; 8(4): 6223. <https://doi.org/10.18282/hrms6223>

COPYRIGHT



Copyright © 2026 by author(s).

Human Resources Management and Services is published by PiscoMed Publishing Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license. <https://creativecommons.org/licenses/by/4.0/>

Abstract: Public-sector performance monitoring depends overwhelmingly on managerial self-report, yet whether such self-assessment reflects the service quality experienced by citizens is rarely tested directly. This study examined human resource capacity in devolved social welfare offices and compared managers' self-assessments against the assessments of the clients they serve. Instruments anchored on identical national administrative standards were administered to the complete population of Local Social Welfare and Development Officers in Pangasinan province, the Philippines ($N = 55$; 35 items, five domains) and to 264 clients drawn by stratified random sampling (21 items, three shared domains). Both scales showed excellent internal consistency (officers, $\alpha = 0.972$; clients, $\alpha = 0.988$). Officers rated overall implementation as Fully Implemented ($M = 3.33$); Human Resource Management and Development was the single lowest-scoring item measured ($M = 3.24$). Neither educational attainment nor years of service predicted implementation in any domain (all $p > 0.05$), indicating that individual credentials do not explain organizational performance. Critically, officers' self-assessments exceeded client assessments on every shared domain (overall gap = 0.55 scale points; $t = 6.63$, $p < 0.001$, Cohen's $d = 0.81$), and the divergence was widest precisely where officer confidence was highest. Structured and open-ended data independently converged on human resource deficits: limited resources, inadequate training, and unmet professional development needs. The findings identify a measurable self-assessment gap in devolved public offices and locate its remedy in organizational human resource systems rather than individual credentialing.

Keywords: human resource capacity; self-other agreement; public sector human resource management; workforce development; competency development; service quality perception; devolved governance

1. Introduction

Human resource capacity has emerged as the decisive constraint on organizational performance in the public sector, particularly where service delivery has been devolved from central agencies to local units. Policy capacity frameworks distinguish the analytical, operational, and political capacities through which a public organization executes its assigned functions (Wu et al., 2015), and the empirical record consistently finds that devolving legal authority does not by itself improve outcomes. Using district-level panel data from decentralized Indonesia, Setiawan et al. (2022) found that these capacities had no significant individual effect on public service delivery but a significant effect when they interacted, meaning a deficit in any one dimension suppresses returns to strength in the others. Malesky and Hutchinson (2016) concluded from a five-country review that Southeast Asian local governments have received devolved responsibilities faster than the capacity to discharge them. Nugroho

and Sujarwoto (2021) reached a similar conclusion for Indonesia, finding that two decades of decentralization produced highly uneven local state capacity, and that development outcomes tracked that capacity more closely than they tracked formal authority. Ghuman and Singh (2013) documented the same pattern across South and Southeast Asia, noting that social welfare functions in particular tend to be devolved on paper well ahead of the capacity to deliver them. Because that capacity resides in people, in their recruitment, training, performance management, and support, devolved service quality is at root a human resource management question (Rondinelli et al., 1983).

The Philippines offers an extended case. The 1991 Local Government Code (Republic Act No. 7160, 1991) devolved social welfare service delivery to local government units (LGUs). The Department of Social Welfare and Development (DSWD) retained authority to set the standards those devolved offices must meet. It does so most concretely through Administrative Order No. 07, series of 2012 (DSWD, 2012), which prescribes standards across five domains: administration and organization, program management, case management, helping interventions, and physical structures and safety. The Magna Carta for Public Social Workers (Republic Act No. 9433, 2007) adds a parallel human resource layer governing compensation, hazard pay, and professional development. Yet Philippine evidence indicates this legal architecture translates unevenly into practice: Capuno (2007) and Canare (2021) both attribute variation in devolved service quality across LGUs to differences in local fiscal capacity and administrative capability rather than to devolution design, and Diokno-Sicat et al. (2020) argue this capacity question has become more consequential as LGUs receive a larger share of national revenue. Within the social welfare sector specifically, Susmerano and Yamada (2022) found that a clearly mandated devolved policy was implemented unevenly depending on the administrative practices of the particular local office responsible, not the clarity of the mandate. Networked and politically embedded governance practices further mediate how devolved authority is exercised at the local level in the Philippines (Porio, 2012). Local Social Welfare and Development Officers (LSWDOs) sit where this architecture meets practice: they head the local Social Welfare and Development (SWD) Office, manage its staff, and are accountable for its outcomes.

Two human resource questions follow, and neither has been examined empirically at the level of the individual devolved office. The first concerns credentials. Public-sector human resource systems are built substantially around credentialing: educational attainment and length of service govern appointment, promotion, and salary grade (Knies et al., 2024; Boyne et al., 1999). Human capital theory holds that such investments raise individual productivity (Becker, 1993; Perry & Wise, 1990). Multilevel accounts qualify this: the emergence of human capital as an organizational resource is a process in which individual attributes do not aggregate straightforwardly into unit-level capability (Ployhart & Moliterno, 2011; Wright & McMahan, 2011). If credentials do not in fact predict how fully an office implements its mandated practices, then human resource investment directed at individual credentialing is misallocated and the binding constraint lies elsewhere in the organizational system.

The second concerns the validity of self-assessment itself. Devolved public-sector performance is monitored predominantly through administrative self-report: offices assess their own compliance against prescribed standards, and those assessments inform oversight, resource allocation, and appraisal. A substantial literature on self-other agreement (SOA) documents that self-ratings systematically exceed the ratings of others in organizational settings (Atwater & Yammarino, 1997; Fleener et al., 2010; Atwater et al., 1998). Meta-analytic evidence places the self-supervisor correlation at modest levels (Harris &

Schaubroeck, 1988), and the self-other correlation across multiple rater sources at similarly limited magnitudes (Conway & Huffcutt, 1997; Heidemeier & Moser, 2009). Mabe and West (1982) further established that self-evaluations of ability relate only weakly to objective criteria. Kruger and Dunning (1999) further demonstrated that individuals least competent in a domain are frequently most confident in it. Meier and O'Toole (2013) warned specifically that subjective, single-source measures of public organizational performance are vulnerable to common-source bias and can generate spurious findings. If these patterns hold in devolved public offices, the instruments on which devolution accountability rests may be systematically overstating service quality, and the human resource development needs they are designed to surface may go undetected.

This study addresses both questions using paired data from a single Philippine province. It surveyed the complete population of LSWDOs in Pangasinan ($N = 55$) on their implementation of DSWD-prescribed practices, and independently surveyed 264 clients of those offices on the same practices. Anchoring both instruments on identical standards allows managerial self-assessment to be compared directly against service-recipient assessment. The study asked: (1) What is the profile of LSWDOs in terms of educational attainment, specialization, and years of service? (2) To what extent are management practices implemented across the five prescribed domains? (3) Does officer profile predict practice implementation? (4) How do clients assess the same practices, and does a significant gap exist between officer and client assessment? (5) What challenges do officers encounter? And (6) what human resource development plan follows from the findings?

This study advances human resource management theory in three specific ways, and it is worth stating at the outset what it does not claim. It does not merely re-confirm that self-assessment is inflated; that proposition is settled (Harris & Schaubroeck, 1988; Conway & Huffcutt, 1997; Heidemeier & Moser, 2009; Mabe & West, 1982). Its contribution lies in identifying the boundary conditions and organizational consequences of that inflation in a setting where the established corrective mechanism is structurally absent.

First, the study relocates self-other agreement from an interpersonal construct to an organizational-systems construct. Within firms, SOA has been theorized principally as a property of individual self-awareness that multisource feedback can correct (Fleenor et al., 2010; Smither et al., 2005; Johnson & Ferstl, 1999). But that entire corrective architecture presupposes an organizational channel through which the other party's rating can reach the focal manager. In the manager-citizen relationship examined here, no such channel exists: clients have no standing in the office's performance management system. SOA in this setting is therefore not a diagnosable individual deficit awaiting feedback but an uncorrected structural feature of the performance measurement regime itself. This reframing carries a direct implication for strategic HRM in public organizations. Models linking HR systems to organizational performance assume that performance information is accurate (Knies et al., 2024; Boyne et al., 1999). If the signal on which HR decisions rest is systematically inflated, the alignment those models presuppose cannot be achieved, however well the HR architecture is otherwise designed.

Second, the study specifies a competency-based mechanism for the gap rather than an attributional one. The dominant explanations for self-assessment inflation are motivational, ego-protection and impression management (Ashford, 1989), or metacognitive (Kruger & Dunning, 1999). The present design allows a third possibility to be examined: that inflation reflects an absent evaluative competency. Program evaluation and human resource management are the two systems-level competencies measured here. If officers weakest in these are also those least

equipped to detect divergence between their own assessments and their clients', then the gap is a competency deficit with a competency-based remedy. On that account, competency modeling frameworks (Campion et al., 2011) become the relevant intervention logic, rather than feedback provision alone. This is a testable proposition about the content of the deficit, not merely its existence.

Third, the study contributes to public workforce development theory by distinguishing credential-based from system-based human capital. Human capital theory treats education and tenure as investments that raise productivity (Becker, 1993), and public appointment systems operationalize that assumption directly. Multilevel accounts, however, hold that individual human capital becomes an organizational resource only through emergence processes that individual attributes do not themselves determine (Ployhart & Moliterno, 2011; Wright & McMahan, 2011). By testing credentials against practice implementation across a complete population, the study provides a direct empirical adjudication between these positions in a public-sector setting, with consequences for how devolved workforce investment should be allocated.

2. Literature review

Human resource capacity and the limits of credentialing. Strategic human resource management research in public organizations emphasizes that context conditions the HRM-performance relationship, and that public-sector HR systems face constraints, statutory appointment rules, salary standardization, and political oversight, that private-sector models do not (Knies et al., 2024; Boyne et al., 1999). Human capital theory (Becker, 1993) treats education and experience as investments that raise individual productivity, and public-sector appointment systems accordingly use them as competence proxies. Yet Ployhart and Moliterno (2011) show that human capital emerges as a unit-level resource only through processes that individual attributes alone do not determine. Wright and McMahan (2011) argue further that the field has too often assumed the aggregation of individual human capital into organizational capability rather than demonstrating it. The Philippine decentralization literature reaches a compatible conclusion from a different direction, attributing variation in devolved service quality to LGU-level fiscal and administrative capacity rather than to personnel attributes (Canare, 2021; Capuno, 2007). Whether officer credentials predict office performance is therefore an open empirical question, and one with direct implications for how public HR investment should be allocated.

Street-level bureaucracy and managerial discretion. LSWDOs occupy the position Lipsky (2010) termed the street-level bureaucrat: public employees whose discretionary decisions, in aggregate, constitute the policy citizens actually experience. Street-level workers operate under chronic resource scarcity and ambiguous goals, and cope through routines and rationing that reshape policy in implementation (Tummers et al., 2015). Supervisory leadership at the frontline shapes how street-level workers exercise that discretion and how they orient toward clients (Keulemans & Groeneveld, 2020), while accountability structures for street-level work remain notoriously weak because the discretionary act is difficult to observe (Hupe & Hill, 2007). A systematic review of the field (Chang & Brewer, 2023) confirms that discretion, coping, and organizational context are the dominant explanatory constructs. Because LSWDOs both exercise street-level discretion and manage the staff who exercise it, they are simultaneously frontline workers and frontline managers, a dual position that makes their self-assessment consequential for organizational monitoring in a way that a purely frontline worker's would not be.

Self-other agreement. The SOA literature provides this study's central theoretical apparatus. Atwater and Yammarino (1997) proposed the foundational typology distinguishing over-estimators, under-estimators, in-agreement/good

estimators, and in-agreement/poor estimators, and demonstrated that the degree of congruence between self-ratings and others' ratings moderates the validity of performance prediction (Atwater & Yammarino, 1992). Atwater et al. (1998) subsequently established that SOA is consequential in its own right for individual and organizational outcomes, and the comprehensive review by Fleenor et al. (2010) concluded that over-estimation in particular signals low self-awareness and is associated with poorer effectiveness. The meta-analytic record supports these claims. Harris and Schaubroeck (1988) found only modest self-supervisor and self-peer convergence. Conway and Huffcutt (1997) found that different rater sources hold systematically different perspectives on the same performance. Heidemeier and Moser (2009) modeled the processes generating self-other divergence across studies. Mabe and West (1982) had earlier established that self-evaluations of ability correlate only weakly with objective criteria, and Ashford (1989) articulated why: self-assessment in organizations is shaped by ego-protective motives and by the scarcity of unambiguous feedback. Kruger and Dunning (1999) added the finding that deficits in a domain simultaneously impair performance and the metacognitive capacity to recognize that impairment, so that the least competent are often the most confident. Erben et al. (2019) demonstrated that SOA in leadership communication is directly associated with subordinates' job satisfaction, establishing that the gap is not merely a measurement artifact but has consequences for the people on the receiving end.

Almost the entire SOA literature, however, compares managers with subordinates, peers, or superiors inside a single organization (Fleenor et al., 2010; Smither et al., 2005; Johnson & Ferstl, 1999). The multisource feedback tradition built on it likewise operates within organizational boundaries (Bracken et al., 2016; DeNisi & Murphy, 2017). What has not been examined is whether the same divergence appears between public managers and the citizens their organizations serve. The practical stakes there are higher. Citizens, unlike subordinates, are the intended beneficiaries of the service, and they have no organizational channel through which to correct the record. Meier and O'Toole (2013) argue that subjective, single-source measures of public organizational performance are systematically vulnerable to common-source bias, and Andrews et al. (2006) document persistent divergence between subjective and objective performance measures in public organizations. The present study tests whether the SOA gap crosses the organizational boundary into the manager-citizen relationship.

Service quality and citizen assessment. The service-quality tradition distinguishes procedural compliance, whether a prescribed action was carried out, from experienced service quality, whether the recipient perceived responsiveness, reliability, and adequacy (Parasuraman et al., 1988). In public services specifically, citizen satisfaction is shaped by expectations as much as by delivered performance (Van Ryzin, 2004; James, 2009), which means that a compliant office can nonetheless be experienced as merely adequate. Citizen assessments moreover shape trust in public servants themselves, not merely satisfaction with a transaction (Van Ryzin, 2011). These are separable constructs, and where public-sector monitoring rests on managerial self-report of compliance, that separability is not a methodological footnote but a structural vulnerability in the accountability system.

Social welfare workforce conditions. The conditions under which social welfare personnel work bear directly on the human resource question. Mor Barak et al. (Mor Barak et al., 2001) established in a foundational meta-analysis that organizational support, supervision, and workload are the principal antecedents of retention and turnover in social work and human services. McFadden et al. (2015) found through systematic review that burnout in child protection social work is driven by organizational factors, control, values, fairness, reward, and workload,

rather than individual resilience alone, and Turley et al. (2022) confirmed that effective interventions for wellbeing and retention are organizational rather than individual in character. Blomberg et al. (2015), Hamama (2012), Savaya (Stanley & Sebastine, 2023), and Stanley and Sebastine (Savaya, 2014) independently document that inadequate supervision, excessive workload, and insufficient organizational support drive job stress and burnout across national contexts. The consistent implication is that the human resource systems inside social welfare offices, not the resilience of the individuals staffing them, determine workforce sustainability and, through it, service quality.

3. Materials and methods

Research design. This study used a descriptive-comparative quantitative design, appropriate for characterizing the current state of a phenomenon and testing differences between independent groups rather than establishing causation (Creswell & Creswell, 2018). The design permitted documentation of officer profiles and practice implementation, testing of association between profile and practice, and direct comparison of officer self-assessment against independent client assessment of the same domains.

Locale and respondents. The study was conducted across the Province of Pangasinan, Region I, Philippines, comprising 44 municipalities and four cities organized into six congressional districts (Philippine Statistics Authority, 2021), each maintaining an SWD Office headed by a designated LSWDO. Two independent samples were drawn. For officers, total population sampling was used: all 55 individuals serving as head, officer-in-charge, or acting head of an SWD Office province-wide participated, yielding complete population coverage rather than a sample estimate. For clients, a population of 773 individuals who had availed of SWD Office services during the study period was identified, and 264 were drawn by stratified random sampling proportionate to district, with sample size determined using Yamane's formula at a 95% confidence level. **Table 1** presents the district distribution of both samples.

Table 1. Distribution of LSWDO and client respondents by congressional district.

District	LSWDOs (f)	LSWDOs (%)	Clients (f)	Clients (%)
District 1	10	18.2	23	8.7
District 2	12	21.8	45	17.0
District 3	6	10.9	46	17.4
District 4	5	9.1	25	9.5
District 5	12	21.8	83	31.4
District 6	10	18.2	42	15.9
Total	55	100.0	264	100.0

Instrumentation. Two structured questionnaires were developed for this study, both anchored directly on the service-delivery standards specified in DSWD Administrative Order No. 07, series of 2012 (DSWD, 2012). The officer instrument comprised 35 items across five domains, Administration and Organization, Program Management, Case Management, Helping Interventions, and Physical Structures and Safety (7 items each), rated from 1 (Not Implemented) to 4 (Fully Implemented). The client instrument comprised 21 items across the three domains clients are positioned to observe directly, Case Management, Helping Interventions, and Physical Structures and Safety, using identical item stems rendered in parallel English and Filipino, rated from 1 (Not Satisfied) to 4 (Highly Satisfied). Anchoring both instruments on the same administrative standards is what permits direct comparison of the two perspectives on the same

underlying practices. This follows the logic of SOA designs, in which self- and other-ratings must reference a common behavioral referent to be comparable (Atwater & Yammarino, 1997; Fleenor et al., 2010). Both instruments were reviewed by subject-matter experts for content clarity, comprehensiveness, and correspondence to the administrative order prior to administration. An open-ended item on challenges encountered was included in the officer instrument.

Instrument development and content validation. Instrument construction followed a four-stage procedure. In the first stage, the operative provisions of DSWD Administrative Order No. 07, series of 2012 (DSWD, 2012) were content-analyzed and decomposed into discrete, observable service-delivery requirements. Because the administrative order is itself the authoritative statement of what these offices are required to do, deriving items directly from its provisions establishes content validity by construction: the item universe is the mandate universe. In the second stage, the extracted requirements were written as behaviorally specific items and allocated to the five domains the order itself defines, yielding a balanced structure of seven items per domain. In the third stage, a panel of subject-matter experts comprising social work academics and senior practitioners reviewed every item. Review criteria were clarity, comprehensiveness, correspondence to the source provision, non-redundancy, and appropriateness of the response scale. Items flagged as ambiguous or as conflating two requirements were revised or split. The panel then confirmed that the retained set covered all operative provisions without omission. In the fourth stage, the client instrument was constructed from those domains a service recipient is positioned to observe directly: Case Management, Helping Interventions, and Physical Structures and Safety. The corresponding item stems were rendered in parallel English and Filipino, with back-translation checking to preserve semantic equivalence across the two versions. Maintaining identical item stems across the officer and client instruments is a design requirement of self-other comparison, which is interpretable only when both raters reference a common behavioral referent (Atwater & Yammarino, 1997; Fleenor et al., 2010).

Reliability and construct validity evidence. Internal consistency was assessed using Cronbach's alpha computed from the collected data (Tavakol & Dennick, 2011). The officer instrument demonstrated excellent reliability at domain and full-scale level (Administration and Organization = 0.943; Program Management = 0.945; Case Management = 0.936; Helping Interventions = 0.983; Physical Structures and Safety = 0.963; full 35-item scale = 0.972). The client instrument likewise demonstrated excellent reliability (Case Management = 0.981; Helping Interventions = 0.986; Physical Structures and Safety = 0.982; full 21-item scale = 0.988). All values substantially exceed the conventional 0.70 threshold (Hair et al., 2018).

Because alpha alone is insufficient evidence of construct validity, two additional analyses were conducted. Corrected item-total correlations were computed for every item against the remainder of its domain. All 35 officer items correlated between 0.665 and 0.978 with their corrected domain totals, and all 21 client items between 0.896 and 0.958. No item fell below the 0.30 retention threshold, and none showed a correlation pattern suggesting misassignment. Inter-domain correlations were then examined as evidence of discriminant validity. Officer domain scores correlated between 0.481 and 0.738, substantially below their respective internal consistency coefficients, indicating that the domains measure related but empirically separable aspects of practice rather than a single undifferentiated construct. The highest inter-domain correlation, between Case Management and Helping Interventions ($r = 0.738$), is theoretically expected given that both concern direct client-facing service delivery. **Table 2** reports the full reliability and validity results.

Table 2. Reliability and construct validity evidence for both instruments.

Instrument / Domain	Items	N	Cronbach's α	Corrected item-total r
Officer instrument				
Administration and Organization	7	55	0.943	0.665–0.878
Program Management	7	55	0.945	0.694–0.867
Case Management	7	55	0.936	0.695–0.878
Helping Interventions	7	55	0.983	0.841–0.978
Physical Structures and Safety	7	55	0.963	0.804–0.935
Full officer scale	35	55	0.972	0.665–0.978
Client instrument				
Case Management	7	264	0.981	0.896–0.941
Helping Interventions	7	264	0.986	0.929–0.958
Physical Structures and Safety	7	264	0.982	0.904–0.948
Full client scale	21	264	0.988	0.896–0.958

Note: Inter-domain correlations for the officer instrument ranged from 0.481 to 0.738, all below their respective internal consistency coefficients, indicating related but empirically separable domains.

Confirmatory factor analysis was not conducted. With 55 officer respondents against 35 items, the case-to-parameter ratio falls far below any conventional threshold for stable factor recovery (Hair et al., 2018), and reporting a fitted model on this sample would convey false precision. The domain structure used here is not an inductively derived latent structure requiring empirical confirmation but the a priori classification prescribed by the administrative order the instrument operationalizes; the appropriate validity evidence is therefore content-based and item-analytic, as reported above. This is acknowledged as a limitation and as a priority for replication at larger scale.

Procedural safeguards against common method bias. Because officer data are self-reported, procedural remedies were implemented in line with established recommendations (Podsakoff et al., 2003). Most importantly, the criterion against which officer self-assessment is evaluated in this study is not another self-report but an independent assessment obtained from a separate respondent population, which structurally eliminates common-source bias from the central comparison (Meier & O'Toole, 2013). Client responses were collected independently of the officers whose offices they assessed, and no client response was made accessible to any officer. Respondent anonymity was assured to reduce evaluation apprehension.

Data analysis. Frequencies and percentages described profiles, specializations, and challenges. Weighted means were computed for practice implementation and client assessment using the standard four-point interval scale (1.00–1.75 = Not Implemented / Not Satisfied; 1.76–2.50 = Partially Implemented / Less Satisfied; 2.51–3.25 = Implemented / Satisfied; 3.26–4.00 = Fully Implemented / Highly Satisfied). Analysis proceeded in three stages. First, bivariate Pearson correlations tested the relationship between officer profile and practice implementation. Second, bivariate tests cannot establish whether individual attributes jointly explain implementation variance. Multivariate ordinary least squares regression models were therefore estimated with four simultaneous predictors: educational attainment, years of service, number of management-relevant trainings attended in the preceding two years, and number of concurrent fields of specialization. The last serves as an indicator of functional workload breadth. Models were estimated for overall implementation and separately for each of the five domains. Multicollinearity

was assessed using variance inflation factors. Third, because the regression predictors are all individual-level attributes, a one-way analysis of variance tested whether congressional district, an organizational and contextual rather than individual variable, was associated with implementation, providing a comparison of individual-level against context-level explanatory power. Welch's independent-samples t-test, which does not assume equal variances or group sizes, tested the difference between officer and client assessment on each shared domain, with Cohen's d reported as the effect size and interpreted against conventional benchmarks (Cohen, 1988). Significance was set at $p < 0.05$. Open-ended responses were coded thematically. Analyses were conducted in Python using pandas, NumPy, SciPy, and statsmodels.

Ethical considerations. Ethical clearance was obtained from the institutional research ethics review board prior to data collection (Approval Code: 2023-0194a-PUBLICADMIN; approved 02 March 2023). Participation was voluntary and informed consent was obtained from all respondents. Responses were kept confidential and are reported only in aggregate.

4. Results

4.1 Profile of respondents

Officers were concentrated in Districts 2 and 5 (12 each, 21.8%), with the fewest in District 4 (5, 9.1%); clients were most heavily represented in District 5 (83, 31.4%), reflecting that district's larger service population (**Table 1**). Social Work is the discipline statutorily mandated for this role (Republic Act No. 4373, 1965). Bachelor's-level Social Work was the dominant qualification: 44 officers (80.0%) held a bachelor's degree as their highest qualification and 11 (20.0%) held a master's degree, while no officer held a completed doctoral degree. Years of service were distributed across bands: 24 officers (43.6%) had served 1–5 years, 11 (20.0%) 6–10 years, and 20 (36.4%) 11 years or more, indicating a workforce combining substantial recent intake with a comparably sized long-tenured core. Children and Youth Welfare was the most common specialization (47 officers, 85.5%), followed by Women and Gender Welfare (44, 80.0%), Community Development (39, 70.9%), Persons with Disabilities Welfare (37, 67.3%), and Elderly Welfare (34, 61.8%). Most officers reported multiple concurrent specializations, indicating broad functional load rather than narrow specialization, a workload pattern the social work workforce literature associates with elevated burnout risk (Mor Barak et al., 2001; McFadden et al., 2015) (**Table 3**).

Table 3. Profile of LSWDO respondents (N = 55).

Variable	Category	f	%
Highest educational attainment	Bachelor's degree	44	80.0
	Master's degree	11	20.0
	Doctoral degree	0	0.0
Years of service as LSWDO	1–5 years	24	43.6
	6–10 years	11	20.0
	11 years and above	20	36.4
Field of specialization (multiple response)	Children and Youth Welfare	47	85.5
	Women and Gender Welfare	44	80.0
	Community Development	39	70.9
	Persons with Disabilities Welfare	37	67.3
	Elderly Welfare	34	61.8

4.2. Extent of implementation of management practices

Table 4 reports implementation across the five prescribed domains. Helping Interventions returned the highest weighted mean ($M = 3.49$, $SD = 0.65$, Fully Implemented), followed by Case Management ($M = 3.37$, $SD = 0.69$) and Administration and Organization ($M = 3.37$, $SD = 0.59$), both also Fully Implemented. Program Management ($M = 3.20$, $SD = 0.71$) and Physical Structures and Safety ($M = 3.20$, $SD = 0.72$) fell into the Implemented band. Overall implementation across all 35 items was 3.33 (Fully Implemented).

Item-level analysis locates the human resource deficit precisely. Within Administration and Organization, an otherwise strong domain, Human Resource Management and Development, covering recruitment, selection, training and development, staff support services, performance appraisal, compensation, and volunteer management, was the single lowest-scoring item in the domain ($M = 3.24$, $SD = 0.67$) and the only one falling below the Fully Implemented threshold. The highest-scoring items in the same domain were formulation of a clear vision, mission, and goals ($M = 3.45$, $SD = 0.57$) and communication of that vision within the organization ($M = 3.40$, $SD = 0.53$). Officers are therefore considerably more effective at articulating organizational direction than at managing the people expected to deliver on it.

Within Program Management, the lowest-scoring item was conducting an annual program review and evaluation workshop with beneficiary, staff, and stakeholder participation ($M = 3.07$, $SD = 0.79$), followed by mainstreaming SWD guidelines into the local development plan ($M = 3.15$, $SD = 0.68$); the highest was ensuring services are consistent with program design and intent ($M = 3.33$, $SD = 0.64$). Within Physical Structures and Safety, provision of appropriate and sufficient office space and fixtures scored lowest ($M = 3.09$), and availability of basic utilities highest ($M = 3.33$). Within Case Management, management of the prescribed caseload ceiling per Registered Social Worker scored lowest ($M = 3.05$), and maintenance of confidential yet accessible case records highest ($M = 3.55$).

The pattern is consistent and diagnostic. The two lowest-scoring items across the entire instrument, human resource management and program evaluation, are both systems-level managerial functions. The highest-scoring items are direct-service and vision-articulation functions. An officer corps trained principally in social work methods rather than in organizational management (Mendoza, 2008) implements direct service at a high level while the internal systems that would sustain and evaluate that service trail.

Table 4. Extent of implementation of management practices by domain (N = 55).

Management Practice Domain	M	SD	Descriptive Equivalent
Helping Interventions	3.49	0.65	Fully Implemented
Case Management	3.37	0.69	Fully Implemented
Administration and Organization	3.37	0.59	Fully Implemented
Program Management	3.20	0.71	Implemented
Physical Structures and Safety	3.20	0.72	Implemented
Overall (35 items)	3.33	—	Fully Implemented

4.3. Officer profile and practice implementation

Table 5 presents the correlation between officer profile and practice implementation. Neither educational attainment nor years of service was significantly associated with implementation in any of the five domains, or with overall implementation (all $p > 0.05$). Coefficients for educational attainment ranged from $r = -0.259$ (Program Management, $p = 0.056$) to

$r = 0.152$ (Helping Interventions, $p = 0.266$), with overall implementation at $r = -0.051$ ($p = 0.713$). Coefficients for years of service ranged from $r = -0.177$ (Helping Interventions, $p = 0.195$) to $r = 0.088$ (Administration and Organization, $p = 0.522$), with overall implementation at $r = -0.111$ ($p = 0.422$). Several coefficients were negative, and the strongest single relationship observed, between educational attainment and Program Management, was negative and approached but did not reach significance.

To test whether individual attributes jointly explain implementation variance where they fail to do so individually, multivariate OLS models were estimated with four simultaneous predictors: educational attainment, years of service, number of management-relevant trainings attended in the preceding two years ($M = 3.29$, $SD = 6.93$), and number of concurrent fields of specialization ($M = 3.65$, $SD = 1.62$). The full model for overall implementation was not significant, $F(4, 50) = 0.58$, $p = 0.677$, and explained 4.5% of variance (adjusted $R^2 = -0.032$). No individual predictor approached significance: educational attainment ($B = -0.0004$, $p = 0.998$), years of service ($B = -0.051$, $p = 0.508$), trainings attended ($B = 0.008$, $p = 0.427$), and specializations ($B = 0.034$, $p = 0.426$). Variance inflation factors ranged from 1.33 to 5.42, indicating no disqualifying multicollinearity. Domain-specific models produced the same pattern; none reached significance, with R^2 ranging from 0.044 (Physical Structures and Safety) to 0.120 (Program Management), the latter the strongest model observed and still non-significant, $F(4, 50) = 1.71$, $p = 0.163$ (Table 6).

By contrast, when congressional district, an organizational and contextual variable rather than an individual attribute, was tested as a predictor of overall implementation, the association approached conventional significance, $F(5, 49) = 2.33$, $p = 0.056$, with district means ranging from 3.05 (District 3) to 3.58 (District 5), a spread of 0.53 scale points. The contrast is instructive: four individual attributes entered simultaneously explained 4.5% of implementation variance and produced a model far from significance, while a single contextual variable approached significance on its own. Because district in this study proxies for a bundle of unmeasured organizational conditions, local fiscal capacity, staffing complements, political environment, and executive support, its associational strength relative to individual credentials should be read as directional evidence rather than as identification of a specific causal mechanism.

This null result for individual attributes is substantively informative rather than a methodological failure. Within a complete population, not a sample, of the officers responsible for devolved social welfare delivery across an entire province, individual credentials, tenure, training exposure, and functional breadth jointly explain none of the variation in how fully those offices implement their mandated practices, while contextual variation between districts accounts for a materially larger share.

Table 5. Relationship between officer profile and extent of practice implementation (Pearson r , $N = 55$).

Practice Domain	Education r	p	Service r	p
Administration and Organization	-0.144	0.294	0.088	0.522
Program Management	-0.259	0.056	-0.068	0.623
Case Management	0.025	0.857	-0.130	0.345
Helping Interventions	0.152	0.266	-0.177	0.195
Physical Structures and Safety	0.006	0.965	-0.128	0.352
Overall implementation	-0.051	0.713	-0.111	0.422

Note: No coefficient reached significance at $p < 0.05$. Education coded 1 = bachelor's, 2 = master's; years of service coded 1 = 1–5 years, 2 = 6–10 years, 3 = 11 years and above.

Table 6. Multivariate regression of individual officer attributes on practice implementation (N = 55).

Predictor / Model	B	SE	t	p	VIF
Educational attainment	−0.0004	0.170	−0.00	0.998	5.31
Years of service	−0.051	0.077	−0.67	0.508	5.42
Trainings attended	0.008	0.010	0.80	0.427	1.33
Fields of specialization	0.034	0.043	0.80	0.426	4.77
Model: overall implementation	R ² = 0.044	Adj. R ² = −0.032	F(4,50) = 0.58	0.677	—
Model: Administration and Org.	R ² = 0.075	Adj. R ² = 0.001	F(4,50) = 1.02	0.409	—
Model: Program Management	R ² = 0.120	Adj. R ² = 0.050	F(4,50) = 1.71	0.163	—
Model: Case Management	R ² = 0.100	Adj. R ² = 0.027	F(4,50) = 1.38	0.254	—
Model: Helping Interventions	R ² = 0.070	Adj. R ² = −0.004	F(4,50) = 0.95	0.445	—
Model: Physical Structures	R ² = 0.044	Adj. R ² = −0.032	F(4,50) = 0.58	0.680	—
Context: district (one-way ANOVA)	—	—	F(5,49) = 2.33	0.056	—

Note: Coefficients shown are from the model predicting overall implementation. No predictor or model reached significance at $p < 0.05$. District, a contextual rather than individual variable, is reported for comparison.

4.4. The self-assessment gap: officer versus client assessment

Table 7 presents the central comparison. On each of the three domains assessed by both officers and clients, officer self-assessment significantly exceeded client assessment, with large effect sizes throughout. For Case Management, officers reported $M = 3.37$ ($SD = 0.58$) against clients' $M = 2.85$ ($SD = 0.72$), a gap of 0.53 scale points ($t = 5.88$, $p < 0.001$, $d = 0.76$). For Helping Interventions, officers reported $M = 3.49$ ($SD = 0.62$) against clients' $M = 2.81$ ($SD = 0.76$), a gap of 0.68 points ($t = 7.07$, $p < 0.001$, $d = 0.91$). For Physical Structures and Safety, officers reported $M = 3.20$ ($SD = 0.65$) against clients' $M = 2.76$ ($SD = 0.76$), a gap of 0.44 points ($t = 4.44$, $p < 0.001$, $d = 0.60$). Across all 21 shared items, officers reported $M = 3.35$ ($SD = 0.52$) against clients' $M = 2.81$ ($SD = 0.70$), an overall gap of 0.55 points ($t = 6.63$, $p < 0.001$, $d = 0.81$).

The gap is not uniform, and its distribution is diagnostic. It is largest precisely in the domain officers rated most highly: Helping Interventions, where officers recorded their strongest self-assessment (3.49, the highest score in the entire officer instrument), produced the widest divergence from client assessment (0.68 points, $d = 0.91$). Conversely, Physical Structures and Safety, among the domains officers rated weakest, produced the narrowest gap (0.44 points, $d = 0.60$). Officer confidence and officer-client divergence move together rather than in opposition.

The practical consequence is a systematic difference in descriptive classification. On every shared domain, officers classified their practice as Fully Implemented or at the upper boundary of Implemented, while clients classified the same practices as Satisfied, never reaching Highly Satisfied on any item. The two groups are not describing the same service.

Table 7. Officer self-assessment versus client assessment on shared practice domains.

Domain	Officer M (SD)	Client M (SD)	Gap	t	d
Case Management	3.37 (0.58)	2.85 (0.72)	0.53	5.88***	0.76
Helping Interventions	3.49 (0.62)	2.81 (0.76)	0.68	7.07***	0.91
Physical Structures and Safety	3.20 (0.65)	2.76 (0.76)	0.44	4.44***	0.60

Continued table:

Domain	Officer M (SD)	Client M (SD)	Gap	t	d
Overall (21 shared items)	3.35 (0.52)	2.81 (0.70)	0.55	6.63***	0.81

Note: Officers $N = 55$; clients $N = 264$. Welch's independent-samples t -test. *** $p < 0.001$. d = Cohen's d .

4.5 Challenges encountered

Table 8 ranks the challenges officers reported. Limited resources, covering budget, staffing, and equipment, was cited by 49 of 55 officers (89.1%), followed by inadequate facilities or infrastructure (21, 38.2%), inadequate training or skills (19, 34.5%), conflicts among staff (7, 12.7%), and lack of support from higher authorities (6, 10.9%). Three of the five, resources, training, and staff conflict, are human resource management functions.

Thematic analysis of open-ended responses surfaced concerns not captured in the structured list, and these converge on the same deficit. Officers reported political intervention in office operations, described as bureaucratic friction and reduced operational autonomy; absence of technical staff to support case management and court-related casework, leaving Registered Social Workers to absorb that workload directly; disruptive transfer and reassignment to offices with substantially different functions; hazard pay perceived as falling below the standard set by the Magna Carta for Public Social Workers (Republic Act No. 9433, 2007); and partial implementation of that statute more generally. Two further themes are directly diagnostic of the human resource development gap identified quantitatively: officers explicitly requested support for continuing professional education, including graduate study, and reported that staff were emotionally unprepared for the demands of client work and required additional training to manage it. That the weakest quantitative item in the Administration and Organization domain, Human Resource Management and Development, is independently corroborated by officers' own unprompted statements is the strongest single piece of evidence in this study for where remediation should be directed. The specific concerns raised, workload, supervision, organizational support, and emotional preparedness, are precisely those the social work workforce literature identifies as the principal drivers of burnout and turnover (Mor Barak et al., 2001; McFadden et al., 2015; Turley et al., 2022; Hamama, 2012).

Table 8. Challenges encountered by LSWDOs in applying management practices ($N = 55$, multiple response).

Challenge	f	%	Rank
Limited resources (budget, staff, equipment)	49	89.1	1
Inadequate facilities or infrastructure	21	38.2	2
Inadequate training or skills	19	34.5	3
Conflicts among staff	7	12.7	4
Lack of support from higher authorities	6	10.9	5

5. Discussion

Credentials do not predict organizational practice. Public-sector appointment, promotion, and compensation systems treat educational attainment and length of service as proxies for organizational competence (Knies et al., 2024; Boyne et al., 1999). In this population they are not, and the null result held across a complete population rather than a sample. The implication is that these systems are allocating on a criterion with no demonstrated relationship to the outcome they are meant to secure. This is consistent with

multilevel accounts of human capital emergence, which hold that individual attributes do not aggregate straightforwardly into unit-level capability (Ployhart & Moliterno, 2011; Wright & McMahan, 2011). It converges as well with the policy capacity literature, which locates devolved performance in the interaction of organizational capacities rather than in the attributes of the post-holder (Wu et al., 2015; Setiawan et al., 2022). Philippine evidence points the same way, finding that service quality varies by LGU capacity rather than by personnel characteristics (Canare, 2021; Capuno, 2007). For practice, credential-raising investment is unlikely to improve service delivery unless matched by investment in the organizational systems officers work within.

The self-assessment gap crosses the organizational boundary. The SOA literature has established over four decades that self-ratings exceed others' ratings within organizations (Harris & Schaubroeck, 1988; Conway & Huffcutt, 1997; Heidemeier & Moser, 2009). It has also shown that over-estimation signals low self-awareness and predicts poorer effectiveness (Fleenor et al., 2010; Atwater et al., 1998), and that the divergence has consequences for those on the receiving end (Erben et al., 2019). What that literature has not examined is whether the same divergence appears outside organizational boundaries. It has developed almost entirely around leader-subordinate, peer, and supervisor dyads (Fleenor et al., 2010; Smither et al., 2005). The present findings show the divergence appearing between public managers and the citizens their organizations exist to serve, at a magnitude comparable to the within-organization effects that literature reports.

The structure of the gap matters more than its size. Divergence was largest in the domain where officers were most confident. This is the pattern Kruger and Dunning (1999) identified, in which deficits in a domain simultaneously impair performance and the metacognitive capacity to detect that impairment. It is also consistent with Ashford's (1989) account of self-assessment as shaped by ego-protective motives and by the scarcity of unambiguous feedback. Officer confidence, in this dataset, is not a signal of service quality; if anything it is a mild counter-indicator. The more consequential difference from within-organization settings is structural. Subordinates can, at least in principle, correct a manager's self-perception through appraisal or feedback systems. Clients have no such standing. The corrective loop that partially disciplines self-assessment inside organizations is absent here entirely, which is why the gap can persist undetected rather than self-correcting over time.

Implications for public-sector performance monitoring. The accountability architecture of devolved social welfare, in the Philippines and across the region, rests substantially on administrative self-report. Meier and O'Toole (2013) warned that subjective, single-source measures of public organizational performance are vulnerable to common-source bias and can generate spurious conclusions. Andrews et al. (2006) documented persistent divergence between subjective and objective performance measures. This study supplies a direct measurement of that divergence in a devolved setting, and shows it is not random: self-assessment overstates quality most where confidence is highest. Monitoring instruments built on self-report are therefore biased toward reporting adequacy, and will fail to surface development needs precisely in the offices least likely to self-identify them. The remedy is not to abandon self-assessment but to pair it with independent service-recipient measurement. Presenting such data is necessary but may not be sufficient: professionals routinely discount external evidence in favour of their own judgement unless granted some control over how it is used (Dietvorst et al., 2018). Multisource feedback research indicates such pairing improves performance when the feedback is credible and actionable (Smither et al., 2005; Johnson & Ferstl, 1999).

The present design demonstrates feasibility: anchoring both instruments on identical standards produces directly comparable data at modest cost.

Locating the human resource deficit. Four independent lines of evidence, reported above, converge on the same conclusion: the weakest practice areas are systems-level managerial functions, not direct-service functions, and officers themselves identify the same gap without prompting. The convergence is what gives the finding weight. Quantitative item scores, structured challenge rankings, and unprompted open-ended responses are methodologically independent, and agreement across all three is unlikely to be an artifact of any one instrument. The conditions officers name are workload, supervision, support, and reward. These are precisely the drivers of burnout and turnover identified in the social work workforce literature (Mor Barak et al., 2001; McFadden et al., 2015), and systematic review shows they respond to organizational rather than individual-level intervention (Turley et al., 2022).

This convergence also suggests a mechanism for the self-assessment gap itself. Officers are trained principally in direct social work methods (Mendoza, 2008), not in human resource management, program evaluation, or performance measurement. They may therefore lack the evaluative apparatus required to detect divergence between their own assessments and their clients' experience. On this reading, the deficit in evaluative capacity and the inflation of self-assessment are one deficit observed from two directions. If so, the remedy lies in explicit, behaviorally anchored development of the evaluative competencies concerned (Campion et al., 2011), not in feedback provision alone.

Positioning against the international public HRM evidence. Knies et al. (2024) argue that the HRM-performance relationship in public organizations is contingent on context to a degree private-sector models understate. The present finding that district-level context outperformed four individual attributes combined is a direct instance of that contingency. Comparative decentralization research reaches the same conclusion from the macro side: Setiawan et al. (2022) found local government capacities significant only in interaction, and Nugroho and Sujarwoto (2021) found Indonesian development outcomes tracking capacity rather than devolved authority. This study supplies the office-level counterpart, showing the same displacement of explanatory power from individual to organizational factors at the smallest unit of devolved administration. Where it departs from that literature is in measurement strategy. Public HRM research relies overwhelmingly on manager-reported performance (Meier & O'Toole, 2013). The divergence documented here raises the possibility that some weak or inconsistent HRM-performance associations in that literature reflect inflation in the criterion rather than weakness in the practices evaluated.

Implications for human resource practice: what local governments can do. Several interventions sit within the authority of an individual LGU and require no national action. First, competency development can be redesigned around the two weakest areas. This means replacing generic management training with behaviorally anchored competency standards (Campion et al., 2011) for program evaluation and human resource management. Such standards would specify observable proficiencies: designing and conducting an annual participatory program review, constructing a staff development plan from a documented training-needs analysis, and conducting structured performance appraisal. Assessment against them can be built into the LGU's existing personnel development framework, and needs-based design of this kind substantially outperforms generic provision (Aguinis & Kraiger, 2009; Salas et al., 2012; Noe, 2020). Second, multisource evaluation can be incorporated at low cost by reusing infrastructure LGUs already operate. Client feedback mechanisms mandated under national anti-red-tape requirements can be

extended from transactional service ratings to the practice domains prescribed by the administrative order. Using the same item stems officers self-assess against makes the two ratings directly comparable within the office's own performance cycle. Third, caseload ceilings and specialization breadth are locally adjustable through staffing deployment, and both are established antecedents of burnout in this workforce (Mor Barak et al., 2001; McFadden et al., 2015). Fourth, onboarding and mentoring protocols for reassigned officers require only administrative issuance.

Implications for human resource practice: what requires national action. Other elements exceed local authority. Statutory compensation and hazard pay under the Magna Carta for Public Social Workers (Republic Act No. 9433, 2007) are set nationally, so the shortfalls officers report cannot be remedied by LGU discretion. Creating plantilla positions for technical and case-management support staff requires Department of Budget and Management authorization. Revising the national qualification standards for the LSWDO position is a Civil Service Commission function, and the present findings suggest those standards may over-weight credentials relative to demonstrated competency. Incorporating independent client assessment into the national monitoring framework, as opposed to voluntary local adoption, requires DSWD issuance. The distinction matters practically: recommendations addressed to the wrong level of government are not actionable, and the local-tier measures can proceed immediately without waiting on national reform.

Implications for theory. Three implications follow. For strategic HRM, models linking HR systems to organizational performance assume a valid performance criterion. That assumption is rarely stated and, where no external corrective channel exists, rarely satisfied (Knies et al., 2024; Meier & O'Toole, 2013). For competency management, the concentration of deficits in evaluative rather than direct-service competencies suggests that competency frameworks for public frontline managers should treat performance measurement literacy as core rather than peripheral (Campion et al., 2011). For public workforce development, explanatory power shifted from individual credentials to organizational context. Investment models premised on aggregating individual human capital (Becker, 1993) therefore require the multilevel qualifications Ployhart and Moliterno (2011) and Wright and McMahan (2011) specify. Devolved workforce strategy should be formulated at the level of the office rather than the post, consistent with strategic HRM scholarship positioning workforce capability rather than compliance as the foundation of sustainable advantage (Nastase et al., 2025).

5.1. Proposed human resource development plan

The following plan is derived from the empirical findings rather than from general prescription. Each element is stated with the specific result that generates it, the mechanism by which it addresses that result, and the level of government able to act. **Table 9** summarizes this mapping.

Element 1: Competency development in program evaluation and human resource management. Empirical basis: these were the two lowest-scoring items measured, annual participatory program review ($M = 3.07$) and Human Resource Management and Development ($M = 3.24$), the latter being the only item in the Administration and Organization domain below the Fully Implemented threshold. Mechanism: behaviorally anchored competency standards specifying observable proficiencies, assessed within the LGU's existing personnel development framework (Campion et al., 2011; Treviño-Elizondo & García-Reyes, 2023). Level: local.

Element 2: Institutionalized client feedback within the performance

cycle. Empirical basis: the officer-client gap of 0.55 scale points overall ($d = 0.81$), widest in the domain officers rated highest. Mechanism: extending existing client feedback mechanisms to cover the same practice domains officers self-assess against, using identical item stems so the two ratings are directly comparable. This supplies the corrective channel whose absence permits the gap to persist (Smither et al., 2005; Johnson & Ferstl, 1999). Level: local, with national standardization desirable.

Element 3: Support for continuing professional education. Empirical basis: no officer in the province held a completed doctoral degree, and officers explicitly requested graduate study support in open-ended responses. Mechanism: scholarship and study-leave arrangements. Level: local for study leave; national for scholarship funding.

Element 4: Augmentation of technical and case-management support staff. Empirical basis: caseload management was the lowest-scoring Case Management item ($M = 3.05$), and officers reported in open-ended responses that Registered Social Workers absorb court-related and administrative casework directly for want of technical staff. Mechanism: creation and filling of support positions, relieving professional staff of non-professional workload; workload is among the most robustly established antecedents of burnout and turnover in this workforce (Mor Barak et al., 2001; McFadden et al., 2015). Level: national, requiring plantilla authorization.

Element 5: Emotional intelligence and client-handling training. Empirical basis: officers reported without prompting that staff were emotionally unprepared for the demands of client work. Mechanism: structured training addressing emotional demands, an organizational rather than individual-resilience intervention, consistent with systematic review evidence that effective wellbeing interventions in this workforce are organizational in character (Turley et al., 2022; Hamama, 2012). Level: local.

Element 6: Standardized transfer and onboarding protocols. Empirical basis: officers described reassignment to offices with substantially different functions as professionally destabilizing. Mechanism: mandatory orientation and mentoring on reassignment, issued administratively. Level: local.

Element 7: Safeguards against political interference. Empirical basis: political intervention in office operations was among the most prominent open-ended themes, described as reducing operational autonomy. Mechanism: written delineation of professional and political decision domains; the street-level bureaucracy literature identifies protection of professional discretion as a persistent constraint on frontline practice (Lipsky, 2010; Hupe & Hill, 2007). Level: local for internal protocol; national for statutory protection.

Element 8: Review of statutory compensation and hazard pay. Empirical basis: officers reported hazard pay below the standard set by the Magna Carta for Public Social Workers, and partial implementation of that statute generally. Mechanism: compliance audit and remediation. Level: national.

The sequencing implied by these findings places Elements 1 and 2 first. They are locally actionable and address the two deficits the data identify most strongly. Element 2 is also self-reinforcing. Once client assessment enters the performance cycle, the gap becomes visible to officers as routine information rather than as a research finding. That visibility is the precondition for targeting the remaining elements accurately over time.

Table 9. Human resource development plan mapped to empirical findings and level of government.

Element	Empirical basis	Mechanism	Level
1. Competency development in program evaluation and HRM	Two lowest items measured: annual participatory review (M = 3.07); HRM and Development (M = 3.24)	Behaviorally anchored competency standards in personnel development framework	Local
2. Institutionalized client feedback	Officer–client gap of 0.55 points (d = 0.81), widest where officers most confident	Extend existing feedback mechanisms to same practice domains and item stems	Local / National
3. Continuing professional education support	No officer held a doctoral degree; graduate study requested in open-ended data	Scholarship and study-leave arrangements	Local / National
4. Technical and case-management staff	Caseload management lowest Case Management item (M = 3.05); RSWs absorb court-related casework	Creation and filling of support plantilla positions	National
5. Emotional intelligence and client-handling training	Officers reported staff emotionally unprepared for client work	Structured organizational training intervention	Local
6. Transfer and onboarding protocols	Reassignment described as professionally destabilizing	Mandatory orientation and mentoring on reassignment	Local
7. Safeguards against political interference	Political intervention prominent in open-ended themes	Written delineation of professional and political decision domains	Local / National
8. Statutory compensation and hazard pay review	Hazard pay reported below Magna Carta standard; partial implementation	Compliance audit and remediation	National

Note: Level indicates the tier of government with authority to act. Elements 1 and 2 are locally actionable and address the deficits most strongly evidenced.

5.2. Limitations and future research

Several limitations bound these findings, and two require particular emphasis because they condition how the central claims should be read.

Unmeasured organizational variables. The conclusion that individual attributes do not predict practice implementation rests on four measured predictors: educational attainment, years of service, training exposure, and functional load. It does not and cannot establish that no individual characteristic matters. More importantly, the study did not measure the organizational variables most likely to account for the implementation variance that individual attributes fail to explain. Staffing complement, office budget and per-capita fiscal capacity, caseload volume, organizational size, leadership style, executive support from the local chief executive, and the intensity of political interference were all outside the measurement design, and several were named by officers themselves as their principal constraints. The district-level analysis reported above is suggestive precisely because district proxies for a bundle of these unmeasured conditions, but a proxy is not a measurement, and the analysis cannot identify which specific organizational factors carry the association. The claim this study supports is therefore the narrower one that credential-based individual attributes are not the operative explanation, not the broader one that any particular organizational factor is. Establishing the latter requires a design that measures organizational conditions directly, which is the most important single extension this line of work requires.

External validity and the boundaries of generalization. All respondents were drawn from one province, and the officer sample, while a complete population within that province, is not representative of Philippine LGUs generally. Pangasinan's fiscal capacity, urbanization profile, and political configuration are specific to it. The findings should therefore be read as characterizing devolved social welfare administration in this provincial context, not as establishing general properties of devolved public organizations. Two implications follow for interpretation. First, the magnitude of the officer-client gap should not be treated as a population parameter for devolved offices

elsewhere. What the study establishes is that a gap of substantial magnitude can exist and persist undetected under managerial self-report, not that it takes this particular value generally. Second, the null finding for credentials is bounded by restricted credential range in this population: 80% of officers held a bachelor's degree as their highest qualification and none held a doctorate. A population with greater credential variance could yield different results, and range restriction is a plausible partial explanation for the null. Multi-region replication within the Philippines would establish whether the phenomenon generalizes. Cross-country replication in comparable devolved systems in Indonesia, Thailand, or Vietnam, where the decentralization capacity literature documents parallel patterns, would additionally permit estimation of how the gap varies with local fiscal and organizational conditions.

Design and measurement limitations. The cross-sectional design establishes association and difference rather than causation. Officer self-report is subject to social desirability bias. This is itself part of the phenomenon under study, but it cannot be separated from genuine assessment error in these data. The central comparison, however, is between two independent respondent populations rather than between two self-reports, which structurally removes common-source bias from that specific test (Podsakoff et al., 2003; Meier & O'Toole, 2013). The officer and client datasets contain no respondent-level linking variable, so it is not possible to associate a given client's assessment with the specific officer whose office served them. The comparison is therefore a group-level comparison between two independent samples from the same province, not a matched-pairs analysis, and individual officers cannot be classified into the over-estimator and under-estimator categories that structure the SOA literature (Atwater & Yammarino, 1997; Atwater et al., 1998). Future research should build the linking identifier into data collection from the outset, permitting matched officer-client analysis, multilevel modeling with officers nested within organizational contexts, and testing of whether individual-level SOA predicts office-level performance as it predicts leader effectiveness within firms (Fleenor et al., 2010). The client sample was restricted to individuals who had already accessed services, excluding those unable or unwilling to access them, whose assessments may differ systematically and, if anything, would likely widen rather than narrow the gap. Finally, the sample size precluded confirmatory factor analysis, so the instrument's validity evidence is content-based and item-analytic rather than model-based (Hair et al., 2018).

6. Conclusion

This study examined human resource capacity in devolved public social welfare offices by assessing the management practices of the complete population of Local Social Welfare and Development Officers across a Philippine province and comparing their self-assessments against those of the clients they serve. Three findings stand out. Officer credentials, both educational attainment and years of service, showed no significant relationship with practice implementation in any domain, indicating that individual credentialing does not predict organizational performance and that human resource investment should be directed at organizational systems instead. Human Resource Management and Development was the single weakest practice item within an otherwise strong administrative domain, a finding independently corroborated by officers' own unprompted accounts of inadequate training, absent technical staff, and unmet professional development needs. And officers' self-assessments significantly exceeded client assessments on every shared domain, by an overall margin of 0.55 scale points with a large effect size, with divergence widest precisely where officer confidence was highest.

Taken together, these findings extend the self-other agreement literature

beyond the organizational boundary within which it has developed. In this provincial context, the divergence between self-assessment and external assessment persists, and is consequential, in the relationship between public managers and the citizens they serve. Whether it does so generally in devolved public organizations is a question this single-province study raises rather than settles, and one that multi-region and cross-country replication should address. For human resource practitioners and public-sector administrators, the practical conclusion supported by these data is that performance monitoring built solely on managerial self-report was, in this setting, biased toward reporting adequacy, and that pairing it with independent service-recipient assessment proved both feasible and informative. Given the contextual specificity of the sample, this is offered as a design principle warranting testing elsewhere rather than as an established general finding. For workforce development, the competencies requiring investment are evaluative and managerial rather than direct-service, and the officers concerned have, in their own words, already asked for them.

Author contributions: Conceptualization, RAD and JCB; methodology, RAD and JCB; software, JCB; validation, RAD and JCB; formal analysis, JCB; investigation, RAD; resources, RAD; data curation, RAD and JCB; writing—original draft preparation, RAD and JCB; writing—review and editing, JCB and RAD; visualization, JCB; supervision, JCB; project administration, RAD. Specifically, RAD designed the study, developed the survey instruments in consultation with the expert panel, secured ethical clearance, coordinated fieldwork across the 48 local government units, administered both surveys, and compiled the raw datasets. JCB designed and executed the statistical analysis, including reliability and item-analytic validity testing, the multivariate regression models, and the between-groups comparisons, constructed the tables, and led the theoretical framing and manuscript revision. Both authors interpreted the findings jointly, drafted and revised the manuscript, and have read and agreed to the published version of the manuscript.

Funding: None.

Ethical approval: This study was reviewed and approved by the institutional research ethics review board (approval number: 2023-0194a-PUBLICADMIN; approved 02 March 2023).

Informed consent statement: Informed consent was obtained from all subjects involved in the study.

Data availability statement: The datasets generated and analyzed during the current study are not publicly available in order to protect respondent confidentiality, but are available from the corresponding author on reasonable request.

Acknowledgments: The authors gratefully acknowledge the Local Social Welfare and Development Officers and clients across the 44 municipalities and four cities of Pangasinan who participated in this study. We further extend our appreciation to the Provincial Government of Pangasinan, to Del Carmen Medical Clinic and Hospital, Inc. for their support, and to the Philippine Association of Social Workers, Inc. (PASWI) for their valuable contributions.

Conflicts of interest: The authors declare no conflicts of interest.

References

1. Aguinis H, Kraiger K. (2009). Benefits of training and development for individuals and teams, organizations, and society. *Annual Review of Psychology*, 60, 451–474. <https://doi.org/10.1146/annurev.psych.60.110707.163505>
2. Andrews R, Boyne GA, Walker RM. (2006). Subjective and objective measures of organizational performance: An empirical exploration. In G. A. Boyne, K. J. Meier, L. J. O'Toole, & R. M. Walker (Eds.), *Public service performance: Perspectives on measurement and management* (pp. 14–34). Cambridge University Press.
3. Ashford SJ. (1989). Self-assessments in organizations: A literature review and integrative model. *Research in Organizational Behavior*, 11, 133–174.

4. Atwater LE, Yammarino FJ. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, 45(1), 141–164. <https://doi.org/10.1111/j.1744-6570.1992.tb00848.x>
5. Atwater LE, Yammarino FJ. (1997). Self-other rating agreement: A review and model. In G. R. Ferris (Ed.), *Research in personnel and human resources management* (Vol. 15, pp. 121–174). JAI Press.
6. Atwater LE, Ostroff C, Yammarino FJ, et al. (1998). Self-other agreement: Does it really matter? *Personnel Psychology*, 51(3), 577–598. <https://doi.org/10.1111/j.1744-6570.1998.tb00252.x>
7. Becker GS. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). University of Chicago Press.
8. Blomberg H, Kallio J, Kroll C, et al. (2015). Job stress among social workers: Determinants and attitude effects in the Nordic countries. *British Journal of Social Work*, 45(7), 2089–2105. <https://doi.org/10.1093/bjsw/bcu038>
9. Boyne G, Jenkins G, Poole M. (1999). Human resource management in the public and private sectors: An empirical comparison. *Public Administration*, 77(2), 407–420. <https://doi.org/10.1111/1467-9299.00160>
10. Bracken DW, Rose DS, Church AH. (2016). The evolution and devolution of 360° feedback. *Industrial and Organizational Psychology*, 9(4), 761–794. <https://doi.org/10.1017/iop.2016.93>
11. Campion MA, Fink AA, Ruggeberg BJ, et al. (2011). Doing competencies well: Best practices in competency modeling. *Personnel Psychology*, 64(1), 225–262. <https://doi.org/10.1111/j.1744-6570.2010.01207.x>
12. Canare T. (2021). Decentralization and welfare: Theory and an empirical analysis using Philippine data. *Public Sector Economics*, 45(1), 69–91. <https://doi.org/10.3326/pse.45.1.4>
13. Capuno JJ. (2007). The quality of local governance and development under decentralization. In A. M. Balisacan & H. Hill (Eds.), *The dynamics of regional development: The Philippines in East Asia* (pp. 204–244). Edward Elgar.
14. Chang A, Brewer GA. (2023). Street-level bureaucracy in public administration: A systematic literature review. *Public Management Review*, 25(11), 2191–2212. <https://doi.org/10.1080/14719037.2022.2065517>
15. Cohen J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
16. Conway JM, Huffcutt AI. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance*, 10(4), 331–360. https://doi.org/10.1207/s15327043hup1004_2
17. Creswell JW, Creswell JD. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
18. DeNisi AS, Murphy KR. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology*, 102(3), 421–433. <https://doi.org/10.1037/apl0000085>
19. Department of Social Welfare and Development. (2012). *Guidelines for local government units' social welfare and development service delivery system* (Administrative Order No. 07, Series of 2012). Republic of the Philippines.
20. Dietvorst BJ, Simmons JP, Massey C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3), 1155–1170. <https://doi.org/10.1287/mnsc.2016.2643>
21. Diokno-Sicat CJ, Castillo AFG, Maddawin RB. (2020). Philippine local government public expenditure review: A survey of national government local government support programs (PIDS Discussion Paper Series No. 2020-48). Philippine Institute for Development Studies.
22. Erben J, Schneider FM, Maier M. (2019). In the ear of the beholder: Self-other agreement in leadership communication and its relationship with subordinates' job satisfaction. *International Journal of Business Communication*, 56(4), 505–526. <https://doi.org/10.1177/2329488416672431>
23. Fleenor JW, Smither JW, Atwater LE, et al. (2010). Self-other rating agreement in leadership: A review. *The Leadership Quarterly*, 21(6), 1005–1034. <https://doi.org/10.1016/j.leaqua.2010.10.006>
24. Ghuman BS, Singh R. (2013). Decentralization and delivery of public services in Asia. *Policy and Society*, 32(1), 7–21. <https://doi.org/10.1016/j.polsoc.2013.02.001>
25. Hair JF, Babin BJ, Anderson RE, et al. (2018). *Multivariate data analysis* (8th ed.). Cengage Learning.
26. Hamama L. (2012). Differences between children's social workers and adults' social workers on sense of burnout, work conditions, and organisational social support. *British Journal of Social Work*, 42(7), 1333–1353. <https://doi.org/10.1093/bjsw/bcr135>
27. Harris MM, Schaubroeck J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41(1), 43–62. <https://doi.org/10.1111/j.1744-6570.1988.tb00631.x>
28. Heidemeier H, Moser K. (2009). Self-other agreement in job performance ratings: A meta-analytic test of a process model. *Journal of Applied Psychology*, 94(2), 353–370. <https://doi.org/10.1037/0021-9010.94.2.353>
29. Hupe P, Hill M. (2007). Street-level bureaucracy and public accountability. *Public Administration*, 85(2), 279–299. <https://doi.org/10.1111/j.1467-9299.2007.00650.x>
30. James O. (2009). Evaluating the expectations disconfirmation and expectations anchoring approaches to citizen satisfaction with local public services. *Journal of Public Administration Research and Theory*, 19(1), 107–123. <https://doi.org/10.1093/jopart/mum034>
31. Johnson JW, Ferstl KL. (1999). The effects of interrater and self-other agreement on performance improvement following upward feedback. *Personnel Psychology*, 52(2), 271–303. <https://doi.org/10.1111/j.1744-6570.1999.tb00162.x>
32. Keulemans S, Groeneveld S. (2020). Supervisory leadership at the frontlines: Street-level discretion, supervisor influence, and street-level bureaucrats' attitude towards clients. *Journal of Public Administration Research and Theory*, 30(2), 307–323. <https://doi.org/10.1093/jopart/muz018>
33. Knies E, Boselie P, Gould-Williams J, et al. (2024). Strategic human resource management and public sector performance: Context matters. *The International Journal of Human Resource Management*, 35(14), 2432–2444. <https://doi.org/10.1080/09585192.2017.1407088>

34. Kruger J, Dunning D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
35. Lipsky M. (2010). *Street-level bureaucracy: Dilemmas of the individual in public services* (30th anniversary expanded ed.). Russell Sage Foundation.
36. Mabe PA, West SG. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), 280–296. <https://doi.org/10.1037/0021-9010.67.3.280>
37. Malesky EJ, Hutchinson FE. (2016). Varieties of disappointment: Why has decentralization not delivered on its promises in Southeast Asia? *Journal of Southeast Asian Economies*, 33(2), 125–138. <https://doi.org/10.1355/ae33-2a>
38. McFadden P, Campbell A, Taylor B. (2015). Resilience and burnout in child protection social work: Individual and organisational themes from a systematic literature review. *British Journal of Social Work*, 45(5), 1546–1563. <https://doi.org/10.1093/bjsw/bct210>
39. Meier KJ, O'Toole LJ. (2013). Subjective organizational performance and measurement error: Common source bias and spurious relationships. *Journal of Public Administration Research and Theory*, 23(2), 429–456. <https://doi.org/10.1093/jopart/mus057>
40. Mendoza TL. (2008). *Social welfare and social work* (3rd ed.). Central Book Supply.
41. Mor Barak ME, Nissly JA, Levin A. (2001). Antecedents to retention and turnover among child welfare, social work, and other human service employees: What can we learn from past research? A review and metanalysis. *Social Service Review*, 75(4), 625–661. <https://doi.org/10.1086/323166>
42. Nastase C, Adomnitei A, Apetri A. (2025). Strategic human resource management in the digital era: Technology, transformation, and sustainable advantage. *Merits*, 5(4), 23. <https://doi.org/10.3390/merits5040023>
43. Noe RA. (2020). *Employee training and development* (8th ed.). McGraw-Hill Education.
44. Nugroho Y, Sujarwoto. (2021). Institutions, outputs and outcomes: Two decades of decentralization and state capacity in Indonesia. *Journal of Southeast Asian Economies*, 38(3), 296–319. <https://doi.org/10.1355/ae38-3b>
45. Parasuraman A, Zeithaml VA, Berry LL. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12–40.
46. Perry JL, Wise LR. (1990). The motivational bases of public service. *Public Administration Review*, 50(3), 367–373. <https://doi.org/10.2307/976618>
47. Philippine Statistics Authority. (2021). 2020 census of population and housing: Pangasinan. Philippine Statistics Authority.
48. Ployhart RE, Moliterno TP. (2011). Emergence of the human capital resource: A multilevel model. *Academy of Management Review*, 36(1), 127–150. <https://doi.org/10.5465/amr.2009.0318>
49. Podsakoff PM, MacKenzie SB, Lee JY, et al. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
50. Porio E. (2012). Decentralisation, power and networked governance practices in Metro Manila. *Space and Polity*, 16(1), 7–27. <https://doi.org/10.1080/13562576.2012.698128>
51. Republic Act No. 4373. (1965). An act to regulate the practice of social work and the operation of social work agencies in the Philippines. Republic of the Philippines.
52. Republic Act No. 7160. (1991). The Local Government Code of 1991. Republic of the Philippines.
53. Republic Act No. 9433. (2007). Magna Carta for public social workers. Republic of the Philippines.
54. Rondinelli DA, Nellis JR, Cheema GS. (1983). Decentralization in developing countries: A review of recent experience (World Bank Staff Working Paper No. 581). World Bank.
55. Salas E, Tannenbaum SI, Kraiger K, et al. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest*, 13(2), 74–101. <https://doi.org/10.1177/1529100612436661>
56. Savaya R. (2014). Social worker burnout in Israel: Contribution of daily stressors identified by social workers. *British Journal of Social Work*, 44(5), 1268–1283. <https://doi.org/10.1093/bjsw/bcs193>
57. Setiawan A, Tjiptoherijanto P, Mahi BR, et al. (2022). The impact of local government capacity on public service delivery: Lessons learned from decentralized Indonesia. *Economies*, 10(12), 323. <https://doi.org/10.3390/economies10120323>
58. Smither JW, London M, Reilly RR. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology*, 58(1), 33–66. https://doi.org/10.1111/j.1744-6570.2005.514_1.x
59. Stanley S, Sebastine AJ. (2023). Work-life balance, social support, and burnout: A quantitative study of social workers. *Journal of Social Work*, 23(6), 1135–1155. <https://doi.org/10.1177/14680173231197930>
60. Susmerano E; Yamada K. (2022). Revisiting auxiliary social services for persons with disability: The Philippines case. *International Journal of Social Welfare*, 31(2), 187–202. <https://doi.org/10.1111/ijsw.12498>
61. Tavakol M; Dennick R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
62. Treviño-Elizondo BL; García-Reyes H. (2023). An employee competency development maturity model for Industry 4.0 adoption. *Sustainability*, 15(14), 11371. <https://doi.org/10.3390/su151411371>
63. Tummers LG, Bekkers V, Vink E, et al. (2015). Coping during public service delivery: A conceptualization and systematic review of the literature. *Journal of Public Administration Research and Theory*, 25(4), 1099–1126. <https://doi.org/10.1093/jopart/muu056>
64. Turley R, Roberts S, Foster C, et al. (2022). Staff wellbeing and retention in children's social work: Systematic review of interventions. *Research on Social Work Practice*, 32(3), 281–309. <https://doi.org/10.1177/10497315211052639>

65. Van Ryzin GG. (2004). Expectations, performance, and citizen satisfaction with urban services. *Journal of Policy Analysis and Management*, 23(3), 433–448. <https://doi.org/10.1002/pam.20020>
66. Van Ryzin GG. (2011). Outcomes, process, and trust of civil servants. *Journal of Public Administration Research and Theory*, 21(4), 745–760. <https://doi.org/10.1093/jopart/muq092>
67. Wright PM, McMahan GC. (2011). Exploring human capital: Putting human back into strategic human resource management. *Human Resource Management Journal*, 21(2), 93–104. <https://doi.org/10.1111/j.1748-8583.2010.00165.x>
68. Wu X, Ramesh M, Howlett M. (2015). Policy capacity: A conceptual framework for understanding policy competences and capabilities. *Policy and Society*, 34(3-4), 165–171. <https://doi.org/10.1016/j.polsoc.2015.09.001>