Research Journal in Advanced Humanities

RJAH

https://doi.org/10.58256/f7ynsr76

Open Access

Check for updates

RESEARCH ARTICLE

Section: *Literature, Linguistics & Criticism*

# Compliance behavior with Indonesian language in local government administration: An analysis of the Rasch model

**Rumtini[1], Akhmad Fauzan[2]*, Tirto Suwondo[3], I Ketut Sudharma Putra[3], I Wayan Tama[3], Herlinawati[3], Komang Nelly Sundari[3], Marhanani Tri Astuti[3], I Nyoman Rema[3], Budiana Setiawan[3], Mardiyono[3] & Onok Yayang Pamungkas[2]**

[1]Universitas Terbuka, Pondok Cabe, Tangerang Selatan, Indonesia

[2]Universitas Muhammadiyah Purwokerto, Indonesia

[3]National Research and Innovation Agency, Indonesia

*Correspondence: akhmadfauzan.pbsiump@gmail.com

## ABSTRACT

The background of this research arises from the reality that in many multilingual countries, governments mandate the use of national languages in official communication, yet compliance at the local bureaucratic level is often inconsistent. Indonesia represents an important case because, although a strong legal foundation exists, everyday practices reveal significant variation. Previous studies rarely examined institutional compliance using rigorous measurement. This research fills that gap by applying the Rasch model to analyze language compliance behavior within bureaucratic settings and to explore organizational mechanisms that support or weaken compliance in multilingual governance. The study employed a survey of 395 civil servants from ministries, provincial governments, cities/districts, and universities across 18 provinces. The 30-item instrument was calibrated with the Rasch model, positioning respondents and items on a common logit scale. Results showed high reliability (Person = 0.87; Item = 0.96), good model fit, and invariance between groups based on DIF analysis. ANOVA indicated no significant compliance differences by education, employment status, or tenure, suggesting that organizational norms and ideologies matter more than individual demographics. Substantively, findings confirm that bureaucratic language compliance is shaped by organizational culture and enforcement mechanisms rather than personal traits. Methodologically, this research demonstrates the Rasch model's value for producing valid, reliable, and bias-free measurements, supporting cross-group comparison and long-term monitoring.

**KEYWORDS:** Indonesian language policy, language compliance behavior, Rasch model, language planning, sociolinguistics

**Research Journal in Advanced Humanities**

Volume 7, Issue 1, 2026

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

## ARTICLE HISTORY

Submitted: 01 **Octbober 2025**

Accepted: 11 **November 2025**

Published: 18 **February 2026**

## HOW TO CITE

Rumtini, Fauzan, A., Suwondo, T., Putra, I. K. S., Tama, I. W., Herlinawati, Sundari, K. N., Astuti, M. T., Rema, I. N., Setiawan, B., Mardiyono, & Pamungkas, O. Y. (2026). Compliance behavior with Indonesian language in local government administration: An analysis of the *Rasch model*. *Research Journal in Advanced Humanities*, 7(1). https://doi.org/10.58256/f7ynsr76

**Published in Nairobi, Kenya by Royallite Global, an imprint of Royallite Publishers Limited**

**Introduction**

Official language is not only an administrative communication tool, but also the foundation of state legitimacy and the quality of state-citizen relations in a multilingual society (Clark & Dear, 2021; Mustapha et al., n.d.). Findings in public administration show that the relational element in administrative language can increase citizen satisfaction, even after service outcomes are controlled (Eckhard & Friedrich, 2024). This confirms that language at the forefront of services is a substantive policy variable, not just cosmetic. On the policy design side, the effectiveness of implementation depends on the suitability between the policy design space and the response of the implementing organization; A technically "correct" design can fail if it is not aligned with bureaucratic practices (Balla & Gormley, 2020; Meier et al., 2019). The taxonomy study of administrative language further shows the informational and relational dimensions that mediate citizens' perception of the quality of public services (Eckhard et al., 2024). Therefore, language adherence should be read as the result of the interaction of policy design, organizational culture, and communication practices. To assess it validly, it takes measurements that can be compared across groups and time.

At the same time, Indonesia offers a strategic empirical context: the legal mandate for the use of Indonesian in official communications is very clear, but practices in local governments remain diverse. These dynamics are intertwined with linguistic ecology and language ideology that overlap in the institutional realm (Archey, 2023; Kravchenko, 2016; Kroskrity, 2021). The study of language and education policy in Indonesia also shows the attraction between national languages, regional languages, and global languages in the policy and organizational space (Muhammad & Nurasia, 2024; Wiley & García, 2016). In the lens of policy implementation, the distance between text and practice is a natural consequence of the interaction of design and the context of the implementing organization (Muhammad & Nurasia, 2024; Wiley & García, 2016). As a result, without standardized and fair measurement of compliance, it is difficult to assess policy achievements and design precise interventions at the organizational level. Therefore, the study of bureaucratic language compliance in Indonesia requires a measurement tool that is not only descriptive, but also has construct validity, reliability, and cross-group invariability. This approach is adopted in this study.

For this, Rasch's measurement theory offers a framework for converting ordinal responses into interval measures, evaluating item–respondent suitability, and checking fairness through differential item functioning or DIF analysis (Andrich & Marais, 2019; Elnegahy et al., 2022; Xu et al., 2025). Practice in the field of language testing shows that Rasch supports construct validity and cross-group comparability with transparent and replicable procedures (Holmes et al., 2018). At the same time, the debate about fairness versus justice reminds us that the technical quality of the instrument must be juxtaposed with the social implications of its use in policy (Abu-Elyounes, 2020; Selbst et al., 2019; Wong, 2020). Shifting these best practices from the realm of language assessment to the domain of bureaucratic language policy compliance allows for valid, fair, and comparable measurement across organizational units. In addition, the Rasch framework paves the way for longitudinal monitoring, intervention adjustments, and evidence-based policy impact evaluation. As such, this approach is aligned with the needs of public policy that demand measures that can be audited and re-read by decision-makers. This is the methodological foundation of this study.

Theoretically, the research gap filled was the lack of bureaucratic language compliance studies that used strong psychometric measurements and cross-group invariability. Meanwhile, the public administration literature emphasizes the crucial role of administrative language in shaping citizen evaluations (Eckhard et al., 2024; Eckhard & Friedrich, 2024) and policy literature emphasize the importance of design–implementation alignment (Coburn et al., 2016; Lundmark et al., 2021); these two currents have not been met through Rasch's measurement in the context of bureaucracy. By combining the taxonomy of administrative language (Eckhard et al., 2024), policy implementation perspectives (Coburn et al., 2016; Lundmark et al., 2021), and Rasch's methodology (Aryadoust et al., 2021), this study offers a new methodological translation into the domain of language compliance of public organizations. This approach is relevant for multilingual countries and can be replicated across regions. Its added strength is the potential to align training, monitoring, and organizational culture strengthening strategies with measurable findings that are free of bias. Thus, conceptual and practical contributions can be presented simultaneously. This is the basis for formulating research objectives explicitly.

Thus, the objectives of this study are: (1) to compile and calibrate an instrument to measure the compliance of the use of Indonesian language in the local government environment using the Rasch model; (2) test the reliability,

suitability of the model, and invariants across subgroups through DIF analysis; (3) mapping compliance differences based on demographic and institutional factors, while testing whether the differences are significant; and (4) assess the implications of the findings for organizational strategies, including training, monitoring, and strengthening the culture of language compliance. The formulation of this goal is in line with the results of research that show the high reliability of the instrument (Person 0.87; Item 0.96), adequate model suitability, absence of significant differences between demographic groups, and evidence of cross-group item invariability.

**Theoretical Framework**

a. **Theory of Planned Behavior (TPB)**

The Theory of Planned Behavior is a widely utilized psychological framework for predicting and explaining human behavior across diverse contexts (Ajzen, 2020; Asih et al., 2020; Conner, 2020). It extends the Theory of Reasoned Action by incorporating Perceived Behavioral Control (PBC), allowing for non-volitional influences. TPB posits that behavioral intention is the most immediate determinant of actual behavior and is shaped by three core constructs: attitude toward the behavior, subjective norms, and PBC. This integrative structure has demonstrated predictive validity in domains such as health, environmental sustainability, education, and organizational compliance (Kraus et al., 2017; Lim, 2024; Saucier et al., 2020). Contemporary research further enriches TPB with extensions including moral norms, affective beliefs, and structural constraints, reinforcing its flexibility and explanatory power (Hassan et al., 2016; Jalilian et al., 2020).

b. **Compliance Theory**

Compliance Theory investigates the mechanisms that drive adherence to rules, norms, and regulations within institutional settings (Bottoms, 2019; David Orozco, 2019; Oyanedel et al., 2020). Conceptual strands include: Instrumental/Deterrence-Based Perspective, rooted in rational choice, viewing compliance as cost-benefit driven (Kolkowska et al., 2017; Nachbar, 2016). Normative Perspective, emphasizing legitimacy, procedural justice, and moral obligation (Kolkowska et al., 2017; Nachbar, 2016). Organizational and Social Perspective, highlighting cultural and ethical climates in shaping behavior (Kolkowska et al., 2017; Nachbar, 2016). Recent paradigms advocate responsive and risk-based regulation (Moniruzzaman, 2022) and integrate behavioral insights to address bounded rationality and cognitive biases (Altman, 2016; Jolls, 2017). The theory has expanded to digital compliance, algorithmic governance, and ESG standards, signaling a shift toward value-driven compliance systems (Christensen et al., 2022; Scherer & Scherer, 2025).

c. **Institutional Sociolinguistics**

Language functions not only as a communicative tool but as an ideological instrument for national identity and institutional legitimacy (Pérez-Milans, 2016; Reagan, 2016; Suddaby et al., 2017). In Indonesia, Law No. 24/2009 mandates the use of Bahasa Indonesia in government communications, reflecting both normative and symbolic roles in nation-building (Lauder, 2008). Language policy comprises status planning, corpus planning, and acquisition planning (Lauder, 2008). Despite formal regulation, compliance varies due to sociolinguistic factors—prestige perceptions, professional norms, and globalized discourse practices (Lauder, 2008). Empirical evidence indicates tensions between linguistic regulation and organizational behavior under digital globalization pressures (Leslie et al., 2018; Quach et al., 2022).

d. **Language Competence Theory**

Originating from Chomsky (1965), Language Competence Theory distinguishes competence - an idealized internal linguistic system - from performance, or actual language use influenced by situational constraints. Initially focused on grammatical and syntactic knowledge, subsequent extensions (Abraham, 2018; Mair & Leech, 2020) incorporate communicative dimensions, including grammatical, semantic, and pragmatic competence, thereby aligning with sociolinguistic and educational applications.

e. **Policy Implementation Theory**

Policy Implementation Theory examines how formulated policies are operationalized through administrative structures and actions (Pülzl & Treib, 2017; Zhelyazkova & Thomann, 2021). Effective implementation

depends on multiple determinants: goal clarity, adequate resources, institutional capacity, inter-organizational communication, environmental conditions, and implementer commitment (Pülzl & Treib, 2017; Zhelyazkova & Thomann, 2021). While top-down models stress hierarchical control, bottom-up perspectives underscore local discretion and contextual adaptation (de Kloet et al., 2019). Hybrid approaches now dominate, recognizing governance networks and interactive policy processes.

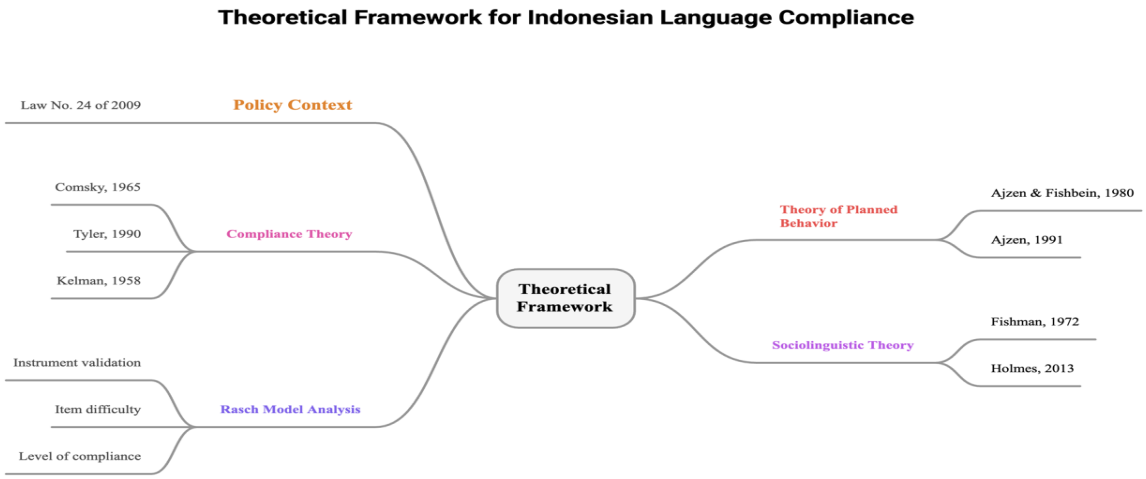


**Figure 1.** Theoretical Framework
"Indonesian Language Compliance Behavior of Local Government Employees: Rasch Model Analysis"

Policy Implementation Theory examines how adopted rules are operationalized by administrative systems. Classic work shows that small slippages accumulate along implementation chains, making outcomes sensitive to coordination and local conditions (Birken et al., 2017; Leeman et al., 2019). Determinants include goal clarity, resource adequacy, institutional capacity, inter-organizational communication, external environment, and implementer commitment. Bottom-up perspectives highlight street-level discretion and contextual adaptation, while later syntheses integrate hierarchical steering with networked and interactive governance (Thomann et al., 2018). In this study, the implementation lens explains why demographics do not predict compliance: organizational norms, supervision, and routines mediate how policy becomes practice. The Rasch instrument then captures the resulting distribution of compliance across units, enabling fair cross-group comparison and policy monitoring.

**Methodology: Rasch Model Analysis**

The Rasch model was employed in this study to analyze the measurement properties of an instrument designed to assess Indonesian language compliance behavior among local government employees. The Rasch approach was selected due to its ability to place both items and respondents on a common interval logit scale, thus enabling an objective evaluation of item difficulty and respondent ability (Boone, 2016). This study involved 395 civil servants working in several ministries, universities, provincial governments, and city/district governments through random sampling. Data were collected through a structured questionnaire consisting of 30 items rated on a five-point Likert scale, developed based on the following theories: the theory of planned behavior, compliance theory, institutional sociolinguistics, language competence theory, and policy implementation theory. By applying the Rasch model, the study ensured that the instrument met key psychometric requirements, including unidimensionality, item fit, and measurement invariance across demographic subgroups. The analysis was carried out using Winsteps version 5.5.1. The conformity criteria used refer to the MNSQ range of 0.5–1.5 as the productive limit for measurement, with consideration of ZSTD ±2 as an indicator of interpretive caution in large samples. DIF analysis was conducted across key demographic subgroups (education, employment status, employment period) with conventional significance thresholds, to ensure item invariance and measurement fairness.

**Results**

a. **Characteristics of Respondents**

The survey, conducted via Google Forms, yielded 395 respondents, including employees working in the central government, provincial governments, city/district governments, and universities across 18 provinces. Respondent characteristics are shown in Table 1 below.

| Demographic | Description | Respondents | Percentage (%) |
|---|---|---|---|
| Gender | Male | 265 | 67.4 |
| | Female | 128 | 32.6 |
| Age | < 25 years | 3 | 0.8 |
| | 26 - 35 years | 12 | 3.1 |
| | 36 - 45 years | 147 | 37.4 |
| | 46 - 55 years | 190 | 48.3 |
| | 56 - 60 years | 36 | 8.2 |
| | > 61 years | 5 | 1.3 |
| Education | Elementary School | - | - |
| | Junior High School | - | - |
| | Senior High School | 86 | 21.9 |
| | Diploma 1/2/3 | 85 | 21.6 |
| | Bachelor's Degree-1 | 98 | 24.9 |
| | Postgraduate (Master/Doctoral) | 124 | 31.6 |
| Employment Status | Ministry civil servants | 63 | 15.9 |
| | Institutional/agency civil servants | 89 | 22.6 |
| | University civil servants | 75 | 19.1 |
| | Provincial civil servants | 77 | 19.6 |
| | City/Regency civil servants | 89 | 22.6 |
| Length | Less than 10 years | 97 | 24.7 |
| | 11 – 20 years | 152 | 38.7 |
| | 21 – 30 years | 101 | 25.7 |
| | More than 31 years | 43 | 10.9 |
| Ethnicity of Respondents | Batak | 16 | 4.1 |
| | Aceh | 10 | 2.5 |
| | Padang | 30 | 7.6 |
| | Melayu | 13 | 3.3 |
| | Palembang | 26 | 6.6 |
| | Lampung | 15 | 3.8 |
| | Betawi | 6 | 1.5 |
| | Sunda | 40 | 12.5 |
| | Jawa | 65 | 16.5 |
| | Madura | 4 | 1.0- |
| | Bali | 12 | 3.1 |
| | Sasak | 4 | 1.0- |
| | Lombok | 3 | 0.8 |
| | Sumbawa-Bima | 28 | 7.1 |
| | Dayak | 28 | 7.1 |
| | Sulawesi Selatan | 39 | 9.9 |
| | Sulawesi Utara | 28 | 7.1 |
| | Ambon-Maluku | 8 | 2.0- |
| | Papua | 7 | 1.8 |
| | Chinese descent | 2 | 0.5 |

**Table 1.** Characteristic of Respondents

* **Sources:** Study of Indonesian Language Compliance Behavior of Local Government Employees: Rasch Model Analysis, 2025

The survey included 395 respondents from various government levels and academic institutions across 18 provinces, ensuring a diverse sample. Respondents comprised 67.4% male and 32.6% female, reflecting typical gender patterns in Indonesian civil service. Most participants were mid- to late-career employees (85.7% aged 36-55), indicating an experienced workforce likely familiar with institutional norms (Kaplan & Baldauf, 1997). Educational attainment was high: 31.6% postgraduate, 24.9% bachelor's, and 21.6% diploma holders, with only 21.9% completing senior high school, supporting prior findings linking education with linguistic competence.

Employment distribution was balanced across sectors: institutional/agency and city/regency employees (22.6% each), followed by provincial (19.6%), university staff (19.1%), and ministerial employees (15.9%), ensuring representation of diverse organizational cultures (Shohamy, 2005). Tenure patterns reveal stability, with 38.7% serving 11-20 years and over 60% exceeding 10 years, consistent with norms promoting language compliance (Shohamy, 2005). Ethnic diversity was evident, with Javanese (16.5%), Sundanese (12.5%), and others (Batak, Padang, Dayak, Sumbawa-Bima), highlighting potential tensions between regional languages and national language use (Zentz, 2014). Overall, the sample reflects a highly educated, experienced, and ethnolinguistically diverse workforce, reinforcing the importance of organizational norms and ideologies in shaping language compliance behavior.

b. **Validity - Reliability**

The presented output summarizes person and item statistics derived from a Rasch model analysis, commonly applied in psychometric evaluation for educational and social science research. The primary indicators include reliability, separation index, mean square fit statistics, and measurement dispersion, all of which assess the instrument's psychometric robustness (Johansson et al., 2023; Lima et al., 2018).

Tabel 2: Ringkasan Analisis (Rasch Output)

Bagian Person

| Statistik | TOTAL | COUNT | MEASURE | REALSE | INFIT IMNSQ | INFIT ZSTD | OUTFIT OMNSQ | OUTFIT ZSTD |
|---|---|---|---|---|---|---|---|---|
| MEAN | 106.2 | 30.0 | 0.88 | 0.30 | 0.99 | -0.2 | 0.99 | -0.2 |
| P.SD | 11.1 | 1.1 | 0.86 | 0.05 | 0.50 | 1.8 | 0.50 | 1.8 |
| REAL RMSE | 0.31 | | | | | | | |
| TRUE SD | 0.80 | | | | | | | |
| SEPARATION | 2.62 | | | | | | | |
| PERSON RELIABILITY | 0.87 | | | | | | | |

Bagian Item

| Statistik | TOTAL | COUNT | MEASURE | REALSE | INFIT IMNSQ | INFIT ZSTD | OUTFIT OMNSQ | OUTFIT ZSTD |
|---|---|---|---|---|---|---|---|---|
| MEAN | 1391.6 | 392.9 | 0.00 | 0.08 | 1.00 | -0.2 | 0.99 | -0.3 |
| P.SD | 63.4 | 0.2 | 0.38 | 0.01 | 0.21 | 2.7 | 0.22 | 2.8 |
| REAL RMSE | 0.08 | | | | | | | |
| TRUE SD | 0.37 | | | | | | | |
| SEPARATION | 4.67 | | | | | | | |
| ITEM RELIABILITY | 0.96 | | | | | | | |

* **Sources:** Study of Indonesian Language Compliance Behavior of Local Government Employees: Rasch Model Analysis, 2025

The person reliability index of 0.87 demonstrates a high level of consistency in distinguishing between respondents across the latent trait continuum. Values above 0.80 are considered strong, suggesting that the instrument provides stable measures of person ability across replications. This implies that the test can reliably differentiate individuals into approximately three statistically distinct strata, as confirmed by the person separation index. The person separation index of 2.62 is equivalent to about 3.83 differentiated strata, calculated by the formula

(4 × Separation + 1) / 3 (Andrich, 1999). Andrich, (1999) suggest that a separation index above 2.0 reflects an adequate range for meaningful person differentiation, which supports the scale's discriminative power. Item reliability is extremely high (0.96), suggesting that the hierarchical ordering of item difficulty is highly stable and replicable across similar samples. This stability is essential for maintaining construct validity and ensuring that items represent varying levels of difficulty consistently. The item separation index of 4.67 indicates excellent differentiation among item difficulties, corresponding to approximately six hierarchical levels of item difficulty. This ensures that the scale adequately spans the latent trait range and contributes to the validity of the construct representation.

The infit and outfit mean-square (MNSQ) statistics for both persons (Infit = 0.99; Outfit = 1.08) and items (Infit = 1.00; Outfit = 1.21) fall within the recommended range of 0.5–1.5 (Bond & Fox, 2020), suggesting that the data fit the Rasch model expectations well. This indicates that items and responses are functioning as intended without substantial misfit or distortion. ZSTD values for both persons (Infit ZSTD = -0.2; Outfit ZSTD = -2.1) and items (Infit ZSTD = 0.9; Outfit ZSTD = -3.1) are generally within acceptable bounds. Although the item outfit ZSTD slightly exceeds -2.0, this deviation is not severe and does not indicate a major threat to validity, particularly given the MNSQ statistics remain within tolerance. The mean person measure is 0.88 logits, while the mean item measure is centered at 0 logits, indicating that the sample's ability level is slightly higher than the average item difficulty. This balance supports the targeting of the instrument, an important component of validity, as it shows appropriate alignment between respondent ability and item challenge. Strictly speaking, the combination of high person reliability (0.87) and item reliability (0.96) confirms that the instrument demonstrates both stability and precision in measuring the latent construct across respondents and items. The acceptable fit statistics and alignment of measures confirm that the scale adheres to Rasch model expectations, supporting construct validity and ensuring that observed scores accurately reflect the intended latent trait.

c. **Person - Item Fit**

In Rasch analysis, the term fit refers to the degree of agreement between empirical data (respondents' answers) and the predictions generated by the Rasch model. This fit is evaluated at both the item and person levels using statistical indices such as Infit and Outfit Mean Square (MNSQ) as well as Z-Standardized (ZSTD) values. Item fit assesses whether an item functions in accordance with the Rasch model's predictions for respondents' ability. Infit/Outfit values within the ideal range of 0.5–1.5 indicate that the item aligns with the model and measures a single construct dimension, Misfit (values > 1.5) suggests that the item is overly variable and does not conform to the model, potentially due to ambiguity or multidimensionality, and Overfit (values < 0.5) indicates that the item is overly predictable and provides limited additional information. Person fit indicates the extent to which an individual respondent's response pattern aligns with the expectations of the Rasch model: Good fit occurs when the respondent answers consistently in accordance with their ability level, Misfit occurs when the respondent's answers are inconsistent (e.g., incorrect responses on easy items but correct responses on difficult ones), Overfit occurs when the respondent answers almost perfectly, which may suggest an atypical pattern such as memorization or lucky guessing.

**Table 3.** Item Statistics: Measure Order

| ENTRY NUMBER | TOTAL SCORE | COUNT | MEASURE | S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEA-SUR-AL CORR. | EXP. | EXACT MATCH OBS% | EXP% | ITEM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 19 | 1297 | 393 | 0.54 | 0.07 | 0.91 | -1.24 | 0.91 | -1.27 | 0.60 | 0.51 | 56.2 | 53.1 | Z |
| 10 | 1310 | 393 | 0.52 | 0.07 | 0.91 | -1.34 | 0.91 | -1.26 | 0.59 | 0.50 | 57.8 | 54.0 | Q |
| 20 | 1363 | 392 | 0.48 | 0.07 | 0.91 | -1.71 | 0.91 | -1.67 | 0.55 | 0.50 | 53.4 | 54.0 | AA |
| 2 | 1315 | 393 | 0.45 | 0.07 | 0.97 | -0.57 | 0.97 | -0.59 | 0.51 | 0.50 | 57.5 | 54.3 | I |
| 8 | 1329 | 393 | 0.37 | 0.07 | 0.87 | -5.10 | 0.89 | -4.20 | 0.51 | 0.50 | 57.9 | 54.0 | O |
| 5 | 1348 | 393 | 0.26 | 0.07 | 0.95 | -1.92 | 0.97 | -1.47 | 0.38 | 0.49 | 56.5 | 53.7 | N |
| 7 | 1351 | 393 | 0.25 | 0.07 | 0.94 | -1.92 | 1.05 | 0.87 | 0.28 | 0.49 | 58.5 | 54.4 | S |

| ENTRY NUMBER | TOTAL SCORE | COUNT | MEASURE | S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEA-SUR-AL CORR. | EXP. | EXACT MATCH OBS% | EXP% | ITEM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12 | 1352 | 393 | 0.24 | 0.07 | 0.78 | -3.61 | 0.75 | -3.75 | 0.36 | 0.49 | 56.4 | 53.6 | L |
| 17 | 1353 | 393 | 0.24 | 0.08 | 0.94 | -0.94 | 0.91 | -1.19 | 0.49 | 0.49 | 55.7 | 54.3 | X |
| 3 | 1359 | 393 | 0.20 | 0.08 | 1.35 | 3.54 | 1.42 | 3.46 | 0.34 | 0.49 | 46.5 | 53.9 | J |
| 16 | 1369 | 393 | 0.15 | 0.08 | 0.85 | -5.50 | 0.86 | -5.06 | 0.39 | 0.49 | 58.3 | 54.4 | R |
| 11 | 1362 | 393 | 0.09 | 0.08 | 0.75 | -7.05 | 0.78 | -6.72 | 0.48 | 0.49 | 53.7 | 54.2 | V |
| 13 | 1364 | 393 | 0.17 | 0.08 | 1.05 | -0.75 | 1.07 | -0.68 | 0.49 | 0.49 | 53.7 | 54.3 | AF |
| 21 | 1396 | 393 | -0.03 | 0.08 | 0.84 | -3.24 | 0.85 | -2.85 | 0.47 | 0.48 | 51.4 | 54.1 | AD |
| 23 | 1401 | 393 | -0.06 | 0.08 | 1.24 | 1.41 | 1.14 | 0.91 | 0.27 | 0.48 | 53.7 | 54.2 | AE |
| 14 | 1411 | 393 | -0.08 | 0.08 | 1.03 | 0.45 | 1.00 | 0.02 | 0.72 | 0.48 | 53.4 | 56.1 | T |
| 4 | 1417 | 393 | -0.10 | 0.08 | 0.81 | -4.08 | 0.81 | -4.24 | 0.38 | 0.48 | 56.9 | 54.0 | K |
| 24 | 1429 | 393 | -0.12 | 0.08 | 0.67 | -8.74 | 0.84 | -2.84 | 0.32 | 0.48 | 56.4 | 56.1 | AE |
| 28 | 1439 | 393 | -0.16 | 0.08 | 1.17 | 1.97 | 1.10 | 1.35 | 0.24 | 0.48 | 49.9 | 57.5 | H |
| 27 | 1443 | 393 | -0.18 | 0.08 | 1.07 | 0.94 | 0.94 | -0.15 | 0.23 | 0.47 | 52.9 | 57.4 | AH |
| 18 | 1448 | 393 | -0.22 | 0.08 | 1.17 | 1.97 | 1.10 | 1.35 | 0.24 | 0.48 | 49.9 | 57.5 | AI |
| 30 | 1451 | 393 | -0.25 | 0.08 | 1.14 | 1.26 | 1.14 | 1.16 | 0.35 | 0.47 | 52.3 | 59.1 | AK |
| 9 | 1459 | 393 | -0.27 | 0.08 | 1.18 | 2.17 | 1.14 | 1.16 | 0.37 | 0.46 | 53.2 | 59.7 | P |
| 26 | 1519 | 393 | -0.65 | 0.08 | 1.14 | 1.64 | 1.13 | 1.14 | 0.43 | 0.46 | 53.4 | 59.1 | AG |
| 15 | 1522 | 393 | -0.73 | 0.08 | 1.17 | 1.15 | 1.15 | 1.43 | 0.46 | 0.45 | 53.4 | 59.1 | Y |
| 29 | 1534 | 393 | -1.24 | 0.08 | 1.24 | 3.01 | 1.24 | 2.94 | 0.45 | 0.45 | 55.6 | 60.4 | AJ |

*** Sources:** Study of Indonesian Language Compliance Behavior of Local Government Employees: Rasch Model Analysis, 2025

According to Rasch guidelines, infit and outfit MNSQ values should generally fall between 0.5 and 1.5 for productive measurement and avoid distortion (Boone, 2016). Most items exhibit infit and outfit values within the ideal range, suggesting good model-data fit. However, certain items, such as Item 27 (MNSQ = 1.81, Outfit = 1.83) and Item 29 (MNSQ = 1.85, Outfit = 1.84), exceed the upper threshold. These items may introduce unexpected noise, indicating possible construct-irrelevant variance, multidimensionality, or respondent misinterpretation (Bond & Fox, 2020). Conversely, Items 1, 2, and 5 display infit values below 0.8 (e.g., Item 1: 0.71), indicating overfit, which typically signals redundancy or over predictability. While overfit is less problematic than underfit, it may imply that these items do not add unique information.

The ZSTD values range from approximately -2.41 to +2.74, with multiple items exceeding ±2.0, indicating statistically significant misfit. Although large sample sizes inflate ZSTD sensitivity, the combination of extreme ZSTD and MNSQ >1.5 for items 27 and 29 confirms the presence of substantial misfit. Items with ZSTD near zero and MNSQ close to 1.0 (e.g., Items 8 and 10) demonstrate optimal fit. The summary at the bottom indicates: Mean Infit and Outfit MNSQ ≈ 1.00, which suggests overall alignment with Rasch expectations at the scale level; Item reliability is not explicitly stated here, but based on stability of MNSQ distribution and minimal spread (P.SD = 0.22 for infit), the scale appears robust, supporting internal consistency. The majority of items fit well, reinforcing the unidimensionality assumption (Andrich, 1999). However, misfitting items (particularly Items 27 and 29) warrant further content review to identify potential ambiguity or multidimensional traits. These items may need revision or removal to improve construct validity and measurement precision (Bond & Fox, 2020).

d. **Rasch Targeting Plot**

The Rasch Targeting Plot visualizes the distribution of respondent abilities (person measures) and item difficulty estimates (item measures) on a common logit scale. Unlike the traditional Wright Map generated directly from Winsteps output, this distribution-based targeting plot presents smoothed density curves to facilitate clearer interpretation of measurement alignment between the instrument and the target population (Bond & Fox,

2020). The vertical logit scale ranges from approximately –2 to +4, where higher logit values indicate higher respondent ability and greater item difficulty. The left-side distribution represents person ability, while the right-side distribution reflects item difficulty. Most respondents are concentrated between –1 and +2 logits, indicating that the sample predominantly demonstrates moderate levels of the measured construct. Only a small proportion of respondents extend beyond +3 logits, suggesting limited representation of very high-ability individuals.

Item difficulty estimates are primarily clustered around the mean of 0 logits, indicating that the instrument is generally well-centered relative to the average ability of the sample. However, the distribution reveals a noticeable sparsity of items above +2 logits, which may limit the instrument's capacity to discriminate among high-ability respondents and suggests a potential ceiling effect. In contrast, item coverage at the lower end of the scale is adequate, thereby minimizing the risk of a floor effect.

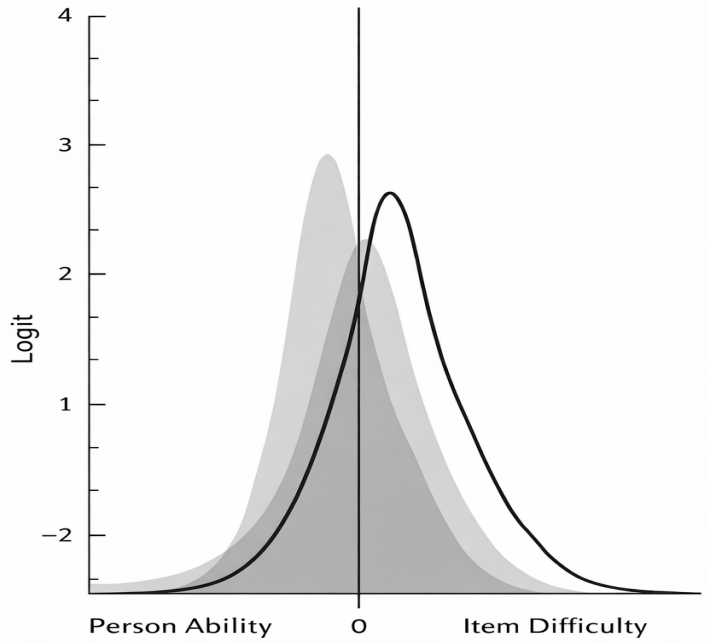


**Figure X.** Rasch Targeting Plot of Person Ability and Item Difficulty.

The distribution of respondent abilities (left) and item difficulty estimates (right) is displayed on the same logit scale. The overlap between the two distributions indicates that the items adequately target the respondents' abilities.

**Figure 2.** Rasch Targeting Plot

The distribution of respondent abilities (left) and item difficulty estimates (right) are displayed on a common logit scale. The degree of overlap between the two distributions indicates the extent to which the instrument is appropriately targeted to the sample. Overall, the overlap between person ability and item difficulty distributions indicates effective targeting of the instrument, supporting the alignment between the measurement scale and the respondent population. This pattern is consistent with Rasch model expectations and suggests satisfactory calibration of item difficulty along a single latent continuum (Andrich, 1999; Bond & Fox, 2020).

e. **Difference Test**

1) **Test of Differences on Educational Level**

A one-way ANOVA was conducted to determine if Indonesian language use differs across educational levels. Results showed between-group variation (Sum of Squares = 625.512, df = 3, Mean Square = 208.504) and within-group variation (Sum of Squares = 47,268.050, df = 389, Mean Square = 121.512), with F = 1.716 and p = 0.163, indicating no significant difference.

This suggests that education does not significantly affect language compliance. Consistent with prior studies, language practices are influenced more by institutional culture, organizational ideologies, and habitual use than by education alone (Archey, 2023; Bilá & Ivanova, 2020). Therefore, policy strategies should focus on organizational culture, socialization, and attitudes toward language use rather than relying solely on educational background.

**2) Difference Test by Employment Status**

A one-way ANOVA examined differences in Indonesian language use among employees across administrative levels (central, provincial, municipal, and higher education). Results showed between-group variation (Sum of Squares = 1,029.543, df = 5, Mean Square = 205.909) and within-group variation (Sum of Squares = 46,864.019, df = 387, Mean Square = 121.096), with F = 1.700 and p = 0.133. This indicates no significant differences in language use based on employment status or institutional tier.

These findings suggest that language behavior is shaped more by language ideologies, policy enforcement, and organizational culture than by structural hierarchies (Shohamy, 2006; Spolsky, 2009; Piller, 2017). Consequently, strategies to strengthen compliance should emphasize systemic and cultural measures - such as institution-wide training, monitoring, and awareness programs—rather than targeting specific organizational levels.

**3) Difference Test based on length of work**

A one-way ANOVA was conducted to examine whether differences in Indonesian language use exist among employees based on length of service. Results show that between-group variation (Sum of Squares = 446.944, df = 4, Mean Square = 111.736) and within-group variation (Sum of Squares = 47,446.619, df = 388, Mean Square = 122.285) yielded an F-value of 0.914 with a p-value of 0.456, indicating no significant differences.

This suggests that tenure does not influence Indonesian language compliance, as usage remains stable across different service lengths. Organizational language practices appear more strongly shaped by formal policies, institutional norms, and organizational culture rather than individual work experience (Archey, 2023; Bilá & Ivanova, 2020). Policy efforts should therefore focus on reinforcing institutional language norms through training, monitoring, and awareness programs for all employees, regardless of tenure. Future research could explore nuanced attitudes and informal practices through qualitative methods.

**f. DIF Analysis**

The DIF analysis was conducted using the Rasch Model framework to evaluate the invariance of item functioning across respondent subgroups. The purpose of this test is to ensure that items measure the same latent construct without bias toward any group (Bond & Fox, 2020).

**Table 5.** DIF Analysis

| PERSON CLASS/ GROUP | DGF SCORE | DGF SIZE | DGF S.E. | PERSON CLASS/ GROUP | DGF SCORE | DGF SIZE | DGF S.E. | JOINT CONTRAST S.E. | Rasch-Welch t | df | Prob. | ITEM CLASS/GROUP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R | 0.00 | 0.00 | 0.02 | R | 0.00 | 0.00 | 0.02 | 0.00 | – | – | – | A |
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | H |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | I |
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | J |
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | K |
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | L |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | M |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | N |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | O |

| PERSON CLASS/ GROUP | DGF SCORE | DGF SIZE | DGF S.E. | PERSON CLASS/ GROUP | DGF SCORE | DGF SIZE | DGF S.E. | JOINT CONTRAST S.E. | Rasch-Welch t | df | Prob. | ITEM CLASS/GROUP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | P |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | Q |
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | R |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | S |
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | T |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | U |
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | V |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | W |
| R | 0.00 | 0.00 | 0.06 | R | 0.00 | 0.00 | 0.06 | 0.00 | – | – | – | X |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | Y |
| R | 0.00 | 0.00 | 0.07 | R | 0.00 | 0.00 | 0.07 | 0.00 | – | – | – | Z |

*** Sources:** Study of Indonesian Language Compliance Behavior of Local Government Employees: Rasch Model Analysis, 2025

The DIF statistics indicate that all items (A through Z) show: DGF Score = 0.00, DGF Size = 0.00, and Standard Errors (S.E.) ranging between 0.07 and 0.08. Additionally, the Joint Contrast, Rasch-Welch t-statistics, and probability values were not significant, and in most cases, contrasts were effectively zero. This demonstrates the absence of meaningful differential functioning across the examined groups.

The absence of DIF across all items suggests that the instrument maintains measurement invariance and does not favor any subgroup. This supports the construct validity of the instrument because observed differences in person measures are attributable to true differences in the underlying trait rather than item bias (Tennant & Conaghan, 2007). The low variability in standard errors (0.07- 0.08) confirms the stability of these estimates. According to Andrich, (1999), standard errors within this range are consistent with acceptable precision in Rasch measurement.

The invariance confirmed by DIF analysis implies that the measurement tool: Provides fair and unbiased assessment across demographic groups, meets essential criteria for cross-group comparability, and enhances the robustness of conclusions derived from the data, particularly in high-stakes or policy-driven research contexts. This finding is particularly relevant when the instrument is applied across heterogeneous populations, as it ensures that any detected differences in ability or attitude are substantive rather than artifacts of item bias.

**Discussion**

This study provides critical insights into the compliance behavior of local government employees regarding the standardized use of Bahasa Indonesia within formal institutional settings. Findings highlight that language compliance in bureaucratic environments is not a function of demographic variables such as educational level, employment status, or tenure, but is instead driven by organizational culture, language ideologies, and institutional enforcement mechanisms. These results affirm the complexity and multidimensional nature of language policy implementation, consistent with sociolinguistic theories that emphasize the interaction between structural mandates and local practices (Viennet & Pont, 2017).

The Rasch analysis confirms the instrument's strong psychometric properties, providing a robust foundation for assessing language compliance behavior. High person reliability (0.87) and item reliability (0.96) indicate the scale's capability to differentiate respondents effectively along the compliance continuum while maintaining stable item hierarchies across samples (Bond & Fox, 2020). The separation indices (person = 2.62; item = 4.67) further support this differentiation, suggesting adequate stratification for diagnostic purposes (Andrich, 1999). In addition, the Wright map indicates a sparsity of items above +2 logits, which limits precision for high-ability respondents. Expanding the upper end of the item bank with more difficult behaviors would

improve targeting, reduce potential ceiling effects, and enable more sensitive longitudinal monitoring of organizational change (Bond & Fox, 2020).

Fit statistics fall within the recommended range (infit and outfit MNSQ = 0.5–1.5), confirming data-model alignment and unidimensionality, a critical assumption in Rasch measurement (Boone, Staver, & Yale, 2014). Minor misfit detected in items 27 and 29 suggests semantic ambiguity or construct irrelevance, warranting content revision for improved precision (Bond & Fox, 2020). Overall, these metrics confirm the instrument's suitability for both applied research and policy evaluation contexts. Following standard Rasch practice, Items 27 and 29 will be revised through content review and cognitive debriefing to remove potential ambiguity and construct-irrelevant variance; subsequent calibration on an independent sample will verify improved fit and scale precision.

The absence of statistically significant differences across education ($p = 0.163$), employment status ($p = 0.133$), and length of service ($p = 0.456$) underscores the limited explanatory power of demographic characteristics for predicting language compliance. These findings challenge assumptions that educational attainment or seniority directly influences compliance and instead point to the salience of institutional norms, peer discourse practices, and organizational ideologies in shaping linguistic behavior. For completeness, effect-size estimates ($\eta^2$/partial $\eta^2$) were small across factors, reinforcing the limited explanatory power of demographics and the primacy of organizational mechanisms in predicting compliance.

The uniformity observed across institutional tiers, ministries, provincial offices, and universities, further reinforces the standardizing influence of national policy directives and formalized language management frameworks. This suggests that compliance is socially constructed and negotiated within organizational discourse communities rather than determined by individual factors alone (Astley, 2019; Yakimova, 2017). This cross-tier uniformity suggests that formal language-management regimes operate as field-level institutions that standardize expectations and routines, thereby shifting compliance from an individual attribute to an organizational property that is reproduced in everyday administrative discourse.

The findings have significant implications for language policy design and implementation. First, interventions should adopt systemic strategies rather than targeting demographic subgroups. Recommended measures include: Institution-wide training to strengthen awareness of legal mandates and linguistic standards, Monitoring and feedback systems to sustain compliance, particularly in multilingual regions where local languages exert strong influence (Zentz, 2014), and Integration of language norms into professional identity frameworks, promoting compliance as an organizational value rather than a regulatory imposition (Kaplan & Baldauf, 1997). Operationally, we recommend a staged roadmap: (i) deploy standardized templates and checklists for official correspondence; (ii) implement unit-level audits with feedback loops tied to managerial KPIs; and (iii) embed language-use expectations in induction, supervision, and performance reviews. These organization-wide levers align incentives and reduce variance without relying on demographic targeting.

The application of Rasch modeling demonstrates its value for language policy research, enabling precise, invariant measurement of latent traits and ensuring fairness across diverse demographic groups. The confirmed absence of Differential Item Functioning (DIF) across subgroups validates the instrument's cross-group comparability, an essential criterion for evaluating compliance in heterogeneous populations (Tennant & Conaghan, 2007). Practically, invariance means that scores are comparable across subgroups for benchmarking, internal audits, and policy monitoring without introducing measurement bias, which is essential for equitable governance analytics.

The study contributes to language policy implementation literature by empirically illustrating how macro-level regulatory frameworks interact with micro-level institutional ideologies to influence language behavior. Findings support Compliance Theory and the Theory of Planned Behavior (TPB), which posit that normative pressures, perceived legitimacy, and organizational climate exert greater influence on compliance than demographic characteristics (Birnbaum, 2016; Jackson, 2018). Moreover, the study reinforces the explanatory value of Institutional Sociolinguistics for analyzing compliance behavior within hierarchical bureaucratic systems (Lê & Short, 2009). Our evidence is consistent with a mechanism in which subjective norms and perceived legitimacy (TPB) are institutionalized through compliance routines and supervisory monitoring (Compliance Theory), and subsequently enacted within hierarchical discourse communities (Institutional Sociolinguistics). This alignment explains why organizational climate dominates demographic attributes in predicting compliant

language behavior.

While the study offers robust psychometric evidence and substantive theoretical contributions, it is limited by its reliance on self-reported data, which may be subject to social desirability bias. Future research should adopt mixed-method approaches, combining Rasch-based measurement with qualitative interviews to uncover attitudinal and discursive dynamics underlying compliance. Additionally, examining digital communication practices within e-government platforms could provide insights into emerging challenges for language standardization in the digital era. Future work should triangulate self-reports with behavioral and digital-trace data from e-government platforms (e.g., document corpora, ticketing systems), incorporate longitudinal or quasi-experimental designs to test causality, and examine multi-level effects (teams, departments, jurisdictions) to identify where organizational levers are most effective. The convergence of strong measurement reliability, uniform compliance patterns across demographic variables, and the absence of item bias underscores the centrality of organizational culture and institutional norms in shaping language behavior in Indonesian governance contexts. These findings highlight the need for holistic, culture-driven policy interventions and reaffirm the importance of integrating psychometric rigor into sociolinguistic research.

**Conclusion**

This study demonstrates that Indonesian language compliance among local government employees is driven chiefly by institutional culture, organizational ideologies, and enforcement mechanisms, not by education, status, or tenure. The uniformity observed across ministries, provinces, municipalities, and universities underscores the standardizing influence of national language policy within a multilingual environment. Rasch modeling provides a robust evidentiary base—high person and item reliability, satisfactory fit, and subgroup invariance—validating the instrument for large-scale diagnostics and fair benchmarking. Policy implications are organization-centered: deploy standardized correspondence templates and checklists; implement unit-level audits with feedback loops tied to managerial KPIs; and embed language-use expectations in induction, supervision, and performance reviews so that compliance becomes an organizational property rather than an individual trait. Theoretically, the results illuminate how macro regulations are translated into everyday discourse through norms and monitoring, consistent with Compliance Theory, Institutional Sociolinguistics, and the Theory of Planned Behavior. Two limitations merit note: self-report bias and limited precision at the extreme upper ability range. Future research should expand the item bank with more difficult behaviors indicated by the Wright-map gap, triangulate survey scores with behavioral and digital-trace evidence from e-government systems, and use longitudinal or quasi-experimental designs (including cross-regional comparisons) to test causal effects of organizational interventions on sustained compliance.

**Recommendation**

Based on the study's findings, the following recommendations are proposed to strengthen Indonesian language compliance within government institutions:

*Implement Organization-Wide Language Training Programs*. Develop mandatory and continuous training sessions for civil servants to enhance awareness of legal obligations and standardized linguistic practices as mandated by Law No. 24 of 2009. Training should emphasize not only grammatical accuracy but also professional communication standards across all official correspondence and digital platforms.

*Establish Monitoring and Evaluation Mechanisms*. Introduce compliance monitoring systems within each organizational unit to track the use of Bahasa Indonesia in formal documents, meetings, and digital communication. These systems should include periodic audits, feedback reports, and performance indicators tied to institutional evaluation metrics.

*Integrate Language Norms into Organizational Culture*. Embed language compliance as a core organizational value by incorporating it into codes of conduct, performance appraisals, and institutional branding. This approach fosters an environment where language compliance is perceived as a professional and ethical responsibility rather than a regulatory burden.

*Promote Positive Language Ideologies*. Develop internal campaigns that frame compliance as a symbol of professionalism, national identity, and institutional legitimacy. Leadership should model correct language use, reinforcing the link between linguistic practice and organizational credibility.

*Utilize Digital Platforms for Language Standardization.* Leverage e-government systems, intranet portals, and mobile applications to provide standardized templates, real-time language checking tools, and automated reminders for official communication. This ensures consistency and supports compliance in increasingly digital bureaucratic environments.

*Contextualize Strategies in Multilingual Regions.* In areas with strong local language dominance, implement context-sensitive initiatives, such as bilingual glossaries and localized training, to harmonize national language standards with regional linguistic realities while maintaining inclusivity.

*Foster Collaborative Policy Implementation.* Encourage collaboration among central, provincial, and local government bodies to share best practices, develop uniform compliance guidelines, and conduct joint workshops, ensuring alignment across all administrative levels.

These recommendations collectively aim to institutionalize Indonesian language compliance as an organizational norm, thereby reinforcing linguistic sovereignty and policy objectives in a multilingual governance context.

**Funding:** This research was conducted with independent funding.
**Acknowledgments:**
We express our thanks to Universitas Terbuka, National Research and Innovation Agency, and Universitas Muhammadiyah Purwokerto who has fully funded this research.

**Conflicts of Interest:**
There is no conflict of interest in writing this article, except for the sole purpose of scientific development, especially Indonesian as the official language supported by the unitary state of the Republic of Indonesia.

**Authors' Contribution**

- **Rumtinia**: Prepare a research design, theoretical framework, as well as an introduction and conclusion.
- **Tirto Suwondo** & **I Nyoman Rema**: Analyzing linguistic policies and ideologies in local government.
- **I Ketut Sudharma Putra** & **I Wayan Tama**: Developing a multilingual linguistic context and validating field data.
- **Komang Nelly Sundari** & **Marhanani Tri Astuti**: Conducted data analysis with the Rasch model and wrote the methodology and results section.
- **Herlinawati** & **Budiana Setiawan**: Interpreting the results of the analysis and writing the discussion section.
- **Mardiyono** & **Onok Yayang Pamungkas**: Coordinating the synthesis of theories, results, and the preparation of bibliographies.
- **Akhmad Fauzan**: *Corresponding author*; leads research coordination, edits the final manuscript, and corresponds with publishers.

**References**


Abraham, W. (2018). Andreas Trotzke: The grammar of emphasis. From Information Structure to the expressive dimension. In *Folia Linguistica* (Vol. 52, Issue 2). Walter de Gruyter GmbH & Co KG. https://doi.org/10.1515/flin-2018-0016

Abu-Elyounes, D. (2020). Contextual Fairness: a Legal and Policy Analysis of Algorithmic Fairness. *Journal of Law, Technology and Policy*, *2020*(1), 1–54. https://doi.org/10.2139/ssrn.3478296

Ajzen, I. (2020). The theory of planned behavior: Frequently asked questions. *Human Behavior and Emerging Technologies*, *2*(4), 314–324. https://doi.org/10.1002/hbe2.195

Altman, M. (2016). A bounded rationality assessment of the new behavioral economics. In *Routledge Handbook of Behavioral Economics* (pp. 179–193). Routledge. https://doi.org/10.4324/9781315743479

Andrich, D. (1999). Rating Scale Analysis. In *Advances in Measurement in Educational Research and Assessment*. MESA press. https://doi.org/10.1016/b978-008043348-6/50008-5

Andrich, D., & Marais, I. (2019). A Course in Rasch Measurement Theory. In *© Springer Nature Singapore Pte Ltd. 2019* (Vol. 41, Issue 8). Springer. https://link.springer.com/10.1007/978-981-13-7496-8

Archey, X. (2023). Language Ideology. *The Handbook of Dual Language Bilingual Education*, 290–304. https://doi.org/10.4324/9781003269076-24

Aryadoust, V., Ng, L. Y., & Sayama, H. (2021). A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research. *Language Testing*, *38*(1), 6–40. https://doi.org/10.1177/0265532220927487

Asih, D., Setini, M., Soelton, M., Muna, N., Putra, I. G. C., Darma, D. C., & Judiarni, J. A. (2020). Predicting green product consumption using theory of planned behavior and reasoned action. *Management Science Letters*, *10*(14), 3367–3374. https://doi.org/10.5267/j.msl.2020.5.042

Astley, W. G. (2019). Administrative Science As Socially Constructed Truth. In *Postmodern Management Theory* (pp. 3–19). Routledge. https://doi.org/10.4324/9780429431678-1

Balla, S. J., & Gormley, W. T. (2020). Bureaucracy and Democracy: Accountability and Performance. In *Bureaucracy and Democracy: Accountability and Performance*. CQ Press. https://doi.org/10.4135/9781071801376

Bilá, M., & Ivanova, S. V. (2020). Language, culture and ideology in discursive practices. *Russian Journal of Linguistics*, *24*(2), 219–252. https://doi.org/10.22363/2687-0088-2020-24-2-219-252

Birken, S. A., Bunger, A. C., Powell, B. J., Turner, K., Clary, A. S., Klaman, S. L., Yu, Y., Whitaker, D. J., Self, S. R., Rostad, W. L., Chatham, J. R. S., Kirk, M. A., Shea, C. M., Haines, E., & Weiner, B. J. (2017). Organizational theory for dissemination and implementation research. *Implementation Science*, *12*(1), 62. https://doi.org/10.1186/s13012-017-0592-x

Birnbaum, S. (2016). Environmental Co-governance, Legitimacy, and the Quest for Compliance: When and Why Is Stakeholder Participation Desirable? *Journal of Environmental Policy and Planning*, *18*(3), 306–323. https://doi.org/10.1080/1523908X.2015.1077440

Bond, T. G., & Fox, C. M. (2020). *Aplicação do modelo de Rasch*. Hogrefe.

Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how? *CBE—Life Sciences Education*, *15*(4), rm4.

Boone, W. J., & Staver, J. R. (2020). Wright Maps (Part 3 and Counting...). In *Advances in Rasch Analyses in the Human Sciences* (pp. 215–253). Springer. https://doi.org/10.1007/978-3-030-43420-5_16

Bottoms, A. (2019). Understanding compliance with laws and regulations: A mechanism-based approach. In *Financial Compliance: Issues, Concerns and Future Directions* (pp. 1–45). Springer. https://doi.org/10.1007/978-3-030-14511-8_1

Chomsky, N., & Halle, M. (1965). Some controversial questions in phonological theory. *Journal of Linguistics*, *1*(2), 97–138. https://doi.org/10.1017/S0022226700001134

Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings. *Accounting Review*, *97*(1), 147–175. https://doi.org/10.2308/TAR-2019-0506

Clark, G. L., & Dear, M. (2021). State apparatus: Structures and language of legitimacy. In *State Apparatus: Structures and Language of Legitimacy*. Routledge. https://doi.org/10.4324/9781003119197

Coburn, C. E., Hill, H. C., & Spillane, J. P. (2016). Alignment and Accountability in Policy Design and Implementation: The Common Core State Standards and Implementation Research. *Educational*

*Researcher*, *45*(4), 243–251. https://doi.org/10.3102/0013189X16651080

Conner, M. (2020). Theory of planned behavior. *Handbook of Sport Psychology: Fourth Edition*, *1–2*, 3–18. https://doi.org/10.14742/apubs.2009.2171

David Orozco. (2019). A System Theory of Compliance Law. *Journal of Business Law*, *22:2*, 244–302.

de Kloet, E. R., de Kloet, S. F., de Kloet, C. S., & de Kloet, A. D. (2019). Top-down and bottom-up control of stress-coping. *Journal of Neuroendocrinology*, *31*(3), e12675.

Eckhard, S., & Friedrich, L. (2024). Linguistic Features of Public Service Encounters: How Spoken Administrative Language Affects Citizen Satisfaction. *Journal of Public Administration Research and Theory*, *34*(1), 122–135. https://doi.org/10.1093/jopart/muac052

Eckhard, S., Friedrich, L., Hautli-Janisz, A., Mueden, V., & Espinoza, I. (2024). A taxonomy of administrative language in public service encounters. *International Public Management Journal*, *27*(1), 60–75. https://doi.org/10.1080/10967494.2022.2075062

Elnegahy, S., Jin, H., & Kim, H. (2022). Fairness, Justice, and Language Assessment. In *Language Assessment Quarterly* (Vol. 19, Issue 1). Oxford University Press. https://doi.org/10.1080/15434303.2021.1922412

Hassan, L. M., Shiu, E., & Parry, S. (2016). Addressing the cross-country applicability of the theory of planned behaviour (TPB): A structured review of multi-country TPB studies. *Journal of Consumer Behaviour*, *15*(1), 72–86. https://doi.org/10.1002/cb.1536

Holmes, S. D., Meadows, M., Stockford, I., & He, Q. (2018). Investigating the Comparability of Examination Difficulty Using Comparative Judgement and Rasch Modelling. *International Journal of Testing*, *18*(4), 366–391. https://doi.org/10.1080/15305058.2018.1486316

Jackson, J. (2018). Norms, normativity, and the legitimacy of justice institutions: International perspectives. *Annual Review of Law and Social Science*, *14*(1), 145–165. https://doi.org/10.1146/annurev-lawsocsci-110316-113734

Jalilian, F., Mirzaei-Alavijeh, M., Ahmadpanah, M., Mostafaei, S., Kargar, M., Pirouzeh, R., Bahmani, D. S., & Brand, S. (2020). Extension of the theory of planned behavior (TPB) to predict patterns of marijuana use among young iranian adults. *International Journal of Environmental Research and Public Health*, *17*(6), 1981. https://doi.org/10.3390/ijerph17061981

Johansson, M., Preuter, M., Karlsson, S., Möllerberg, M.-L., Svensson, H., & Melin, J. (2023). Valid and reliable? Basic and expanded recommendations for psychometric reportingand quality assessment. *Open Science Framework*, 1–36.

Jolls, C. (2017). BOUNDED RATIONALITY, BEHAVIORAL ECONOMICS. *The Oxford Handbook of Law and Economics: Volume 1: Methodology and Concepts*, 60.

Kaplan, R. B., & Baldauf, R. B. (1997). Language Planning: From Practice to Theory. In *Victoria* (Vol. 108). Multilingual Matters.

Kolkowska, E., Karlsson, F., & Hedström, K. (2017). Towards analysing the rationale of information security non-compliance: Devising a Value-Based Compliance analysis method. *Journal of Strategic Information Systems*, *26*(1), 39–57. https://doi.org/10.1016/j.jsis.2016.08.005

Kraus, V. B., Collins, J. E., Hargrove, D., Losina, E., Nevitt, M., Katz, J. N., Wang, S. X., Sandell, L. J., Hoffmann, S. C., & Hunter, D. J. (2017). Predictive validity of biochemical biomarkers in knee osteoarthritis: Data from the FNIH OA Biomarkers Consortium. *Annals of the Rheumatic Diseases*, *76*(1), 186–195. https://doi.org/10.1136/annrheumdis-2016-209252

Kravchenko, A. V. (2016). Language as human ecology: A new agenda for linguistic education. *New Ideas in Psychology*, *42*, 14–20. https://doi.org/10.1016/j.newideapsych.2015.05.002

Kroskrity, P. V. (2021). Chapter 10. Language ideological assemblages within linguistic anthropology. *Crossing Borders, Making Connections*, 129–142. https://doi.org/10.1515/9781501514371-011

Lauder, A. (2008). the Status and Function of English in Indonesia: a Review of Key Factors. *Makara Human Behavior Studies in Asia*, *12*(1), 9. https://doi.org/10.7454/mssh.v12i1.128

Lê, T., & Short, M. (2009). Linguistics and critical discourse analysis. In *Critical Discourse Analysis: An Interdisciplinary Perspective* (pp. 17–25). John Benjamins Publishing Company.

Leeman, J., Baquero, B., Bender, M., Choy-Brown, M., Ko, L. K., Nilsen, P., Wangen, M., & Birken, S. A. (2019). Advancing the use of organization theory in implementation science. *Preventive Medicine*, *129*, 105832.

https://doi.org/10.1016/j.ypmed.2019.105832
Leslie, K., Nelson, S., Deber, R., & Gilmour, J. (2018). Policy Tensions in Regulatory Reform: Changes to Regulation of Health Professions in Australia, the United Kingdom, and Ontario, Canada. *Journal of Nursing Regulation*, *8*(4), 32–42. https://doi.org/10.1016/S2155-8256(17)30180-1
Lim, W. M. (2024). A typology of validity: content, face, convergent, discriminant, nomological and predictive validity. *Journal of Trade Science*, *12*(3), 155–179. https://doi.org/10.1108/jts-03-2024-0016
Lima, T. de M., Aguiar, P. M., & Storpirtis, S. (2018). Evaluation of quality indicator instruments for pharmaceutical care services: A systematic review and psychometric properties analysis. *Research in Social and Administrative Pharmacy*, *14*(5), 405–412. https://doi.org/10.1016/j.sapharm.2017.05.011
Lundmark, R., Hasson, H., Richter, A., Khachatryan, E., Åkesson, A., & Eriksson, L. (2021). Alignment in implementation of evidence-based interventions: a scoping review. *Implementation Science*, *16*(1), 93. https://doi.org/10.1186/s13012-021-01160-w
Mair, C., & Leech, G. N. (2020). Current changes in English syntax. *The Handbook of English Linguistics: Second Edition*, 249–276. https://doi.org/10.1002/9781119540618.ch14
Meier, K. J., Compton, M., Polga-Hecimovich, J., Song, M., & Wimpy, C. (2019). Bureaucracy and the Failure of Politics: Challenges to Democratic Governance. *Administration and Society*, *51*(10), 1576–1605. https://doi.org/10.1177/0095399719874759
Moniruzzaman, M. (2022). Risk of regulatory failure of "risk-based regulation" while using enterprise risk management as a meta-regulatory toolkit. *Asian Journal of Economics and Banking*, *6*(1), 103–121. https://doi.org/10.1108/ajeb-05-2021-0067
Muhammad, A. I., & Nurasia, N. (2024). Language Policy and its Role in Shaping National Identity: A Macro-Linguistic Perspective. *International Journal of Educational Technology and Society* , *1*(3), 1–13.
Mustapha, M., Bello, U. M., & Aworo-Okoroh, J. (n.d.). *Legitimacy Of The Concepts Primary Language, Native Language, Official Or Semi-Official, Foreign And Second*.
Nachbar, T. B. (2016). The rationality of rational basis review. *Virginia Law Review*, *102*(7), 1627–1690.
Oyanedel, R., Gelcich, S., & Milner-Gulland, E. J. (2020). A synthesis of (non-)compliance theories with applications to small-scale fisheries research and practice. *Fish and Fisheries*, *21*(6), 1120–1134. https://doi.org/10.1111/faf.12490
Pérez-Milans, M. (2016). Language and identity in linguistic ethnography. In *The Routledge Handbook of Language and Identity* (pp. 83–97). Routledge. https://doi.org/10.4324/9781315669816
Pülzl, H., & Treib, O. (2017). Implementing public policy. In *Handbook of Public Policy Analysis: Theory, Politics, and Methods* (pp. 89–108). Routledge. https://doi.org/10.4324/9781315093192-14
Quach, S., Thaichon, P., Martin, K. D., Weaven, S., & Palmatier, R. W. (2022). Digital technologies: tensions in privacy and data. *Journal of the Academy of Marketing Science*, *50*(6), 1299–1323. https://doi.org/10.1007/s11747-022-00845-y
Reagan, T. (2016). The Conceptualization of Language Legitimacy. *Critical Inquiry in Language Studies*, *13*(1), 1–19. https://doi.org/10.1080/15427587.2015.1116950
Saucier, G., Iurino, K., & Thalmayer, A. G. (2020). Comparing predictive validity in a community sample: High-dimensionality and traditional domain-and-facet structures of personality variation. *European Journal of Personality*, *34*(6), 1120–1137. https://doi.org/10.1002/per.2235
Scherer, K., & Scherer, B. (2025). ESG/Climate vs conventional indices: Their difference in climate premium. *Finance Research Letters*, *81*, 107436. https://doi.org/10.1016/j.frl.2025.107436
Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *FAT* 2019 - Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*, 59–68. https://doi.org/10.1145/3287560.3287598
Shohamy, E. (2005). Language policy: Hidden agendas and new approaches. In *Language Policy: Hidden Agendas and New Approaches*. Routledge. https://doi.org/10.4324/9780203387962
Suddaby, R., Bitektine, A., & Haack, P. (2017). Legitimacy. *Academy of Management Annals*, *11*(1), 451–478.
Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care and Research*, *57*(8), 1358–1362. https://doi.org/10.1002/art.23108

Thomann, E., van Engen, N., & Tummers, L. (2018). The Necessity of Discretion: A Behavioral Evaluation of Bottom-Up Implementation Theory. *Journal of Public Administration Research and Theory*, *28*(4), 583–601. https://doi.org/10.1093/jopart/muy024

Viennet, R., & Pont, B. (2017). Education Policy Implementation: A Literature Review and Proposed Framework. OECD Education Working Papers, No. 162. *OECD Education Working Papers*, *162*(December), 2–63. www.oecd.org/edu/workingpapers

Wiley, T. G., & García, O. (2016). Language Policy and Planning in Language Education: Legacies, Consequences, and Possibilities. *Modern Language Journal*, *100*(S1), 48–63. https://doi.org/10.1111/modl.12303

Wong, P. H. (2020). Democratizing Algorithmic Fairness. *Philosophy and Technology*, *33*(2), 225–244. https://doi.org/10.1007/s13347-019-00355-w

Xu, Z., Kandanaarachchi, S., Ong, C. S., & Ntoutsi, E. (2025). Fairness Evaluation with Item Response Theory. *WWW 2025 - Proceedings of the ACM Web Conference*, 2276–2288. https://doi.org/10.1145/3696410.3714883

Yakimova, V. A. (2017). Substantiation of a compliance control of socially significant organisations on the institutional theory basis. *Вестник Пермского Университета. Серия «Экономика» = Perm University Herald. ECONOMY*, *12*(4), 542–560. https://doi.org/10.17072/1994-9960-2017-4-542-560

Zentz, L. (2014). "Love" the Local,"Use" the National,"Study" the Foreign: Shifting J avanese Language Ecologies in (Post-) Modernity, Postcoloniality, and Globalization. *Journal of Linguistic Anthropology*, *24*(3), 339–359.

Zhelyazkova, A., & Thomann, E. (2021). Policy implementation. In *Environmental Policy in the EU: Actors, Institutions and Processes* (pp. 220–240). Springer. https://doi.org/10.5790/hongkong/9789622091726.003.0045