

A Diagnostic Tool for Tracking Purpose Practice Over Time

Authors

Abiodun Okunola, Abi Cit, Billy Elly

Date; 24 October 2026

Abstract

This study addresses the absence of a validated, longitudinal instrument for tracking purpose–practice decoupling by developing and testing the Purpose–Practice Decoupling Tracker (PPDT), a composite index integrating disclosure semantics, resource allocation signals, and operational outcome indicators. Drawing on a retrospective panel of 184 publicly listed firms across three reporting periods (2021–2023), the study combines natural language processing of purpose statements with financial and operational ratio analysis. The PPDT achieves 89.4% classification accuracy ($F1 = 0.87$, $AUC = 0.91$) in distinguishing persistently decoupled firms from those demonstrating tightening coupling over time, outperforming static ESG-gap baselines by 23.6 percentage points. Feature importance analysis identifies narrative specificity decay (weight = 0.31) and resource–rhetoric ratio divergence (weight = 0.27) as the strongest temporal predictors. The findings establish decoupling as a measurable trajectory rather than a binary state, offering administrators and regulators a replicable monitoring instrument while extending performative theories of corporate purpose into the domain of longitudinal diagnostic assessment.

Keywords: diagnostic instrument, longitudinal measurement, corporate purpose, performative governance

1. Introduction

1.1 Background

The proliferation of corporate purpose statements over the past decade has generated a parallel proliferation of skepticism regarding their substantive value. Firms increasingly publish commitments to stakeholder welfare, environmental stewardship, and societal contribution, often embedded within sustainability reports and annual disclosures that invoke frameworks such as the Sustainable Development Goals. Yet a substantial body of organizational research demonstrates that such commitments frequently fail to translate into operational reality. Bromley and Powell [1] distinguish between policy–practice decoupling, wherein a policy is adopted but not implemented, and means–ends decoupling, wherein implementation occurs but fails to produce intended outcomes. This distinction is consequential: it suggests that even firms that appear to “do something” may remain decoupled at a deeper level.

Recent scholarship has extended this concern into the specific domain of corporate purpose. Okeke [2] develops a performative theory of corporate purpose, arguing that purpose communication functions as a governance mechanism whose credibility depends on patterned discursive acts that bind firms to evaluative scrutiny. This performative lens implies that purpose claims are not merely descriptive but constitutive: they create expectations that can be subsequently validated or violated through observable practice. The diagnostic challenge, therefore, is not simply whether a firm has a purpose statement, but whether the trajectory of its practices remains coherent with the commitments that statement performs.

1.2 Problem Statement

Existing approaches to detecting decoupling suffer from three interrelated limitations. First, most diagnostic instruments are cross-sectional: they assess the gap between claim and practice at a single point in time, treating decoupling as a state rather than a process [1][3]. This is theoretically inadequate given evidence that decoupling emerges, intensifies, or attenuates over multiple reporting cycles. Second, prevailing tools rely predominantly on disclosure-based indicators, which are susceptible to strategic manipulation; firms learn to craft disclosures that satisfy external auditors while leaving operational practices untouched. Okeke [4] demonstrates that corporate environmental misrepresentation requires multidimensional assessment to detect, yet no equivalent instrument exists for purpose–practice decoupling specifically. Third, the few longitudinal studies that exist employ qualitative case designs that cannot be scaled or standardized, leaving practitioners without a replicable monitoring tool.

The unsolved issue, therefore, is the absence of a validated, quantitative instrument capable of tracking purpose–practice decoupling as a temporal trajectory across firms, sectors, and reporting periods, using indicators that resist purely discursive manipulation.

1.3 Objectives of the Study

General objective: To develop and validate a diagnostic instrument for tracking purpose–practice decoupling over time.

Specific objectives:

- To identify the combination of disclosure, resource allocation, and operational indicators that most accurately predicts persistent purpose–practice decoupling across reporting periods.
- To design a composite index (the PPDT) that captures decoupling as a trajectory rather than a static gap.
- To validate the instrument against held-out firm-year observations and compare its performance to baseline static methods.

1.4 Research Questions

1. What combination of linguistic, resource, and operational variables most accurately distinguishes persistently decoupled firms from those exhibiting tightening coupling over time?
2. How does the PPDT compare to static disclosure-gap methods in terms of classification accuracy and temporal lead time?
3. What are the practical implementation requirements for deploying such an instrument in regulatory or administrative monitoring contexts?

1.5 Significance of the Study

For practitioners and administrators, the PPDT offers a monitoring mechanism that can flag deteriorating purpose–practice alignment before reputational or regulatory crises materialize. For policymakers, the instrument provides an evidence base for proportionality in oversight: firms demonstrating tightening coupling may warrant lighter-touch monitoring, while persistently decoupled firms may require escalated scrutiny. For academic literature, the study contributes a quantitative operationalization of decoupling as trajectory, extending Bromley and Powell’s [1] conceptual distinction into measurable form and providing empirical support for Okeke’s [2] performative framework by demonstrating that discursive patterns have predictive value when combined with material indicators. For future researchers, the instrument provides a validated foundation for comparative studies across governance contexts, sectors, and regulatory regimes.

1.6 Scope and Limitations

The study is bounded by three parameters. Temporally, it covers the 2021–2023 reporting periods, a window sufficient to observe trajectory variation but insufficient for assessing long-run decoupling dynamics. Geographically, the sample comprises publicly listed firms in

developed Anglophone markets, limiting generalizability to emerging or state-directed governance contexts. Substantively, the instrument focuses on purpose–practice decoupling as defined through publicly observable indicators; it does not access internal managerial decision processes. Key limitations include reliance on simulated data for certain operational indicators where public disclosure is incomplete, and the assumption that historical patterns of decoupling trajectory stability will persist into future periods.

2. Literature Review

2.1 Conceptual Review

Purpose–practice decoupling refers to the divergence between an organization’s articulated purpose commitments and its observable operational practices. Building on Bromley and Powell’s [1] foundational distinction, this study treats decoupling as existing on a continuum rather than as a binary category. At one pole, firms demonstrate tightening coupling: purpose claims become increasingly specific, resource allocations align with stated priorities, and operational outcomes corroborate commitments. At the other pole, firms exhibit persistent decoupling: rhetorical commitments remain stable or expand while resource allocation and operational indicators stagnate or diverge.

Corporate purpose is understood here not as a singular statement but as a patterned set of commitments that performatively constitute organizational obligations [2]. This performative dimension is critical for measurement: purpose claims create evaluative expectations that can be tracked against subsequent observable behavior. The credibility of purpose communication depends on what Okeke [2] terms “discursive costliness”—the reputational exposure created by specific, affect-laden, and temporally sustained commitments that are harder to fabricate than generic language.

Diagnostic instruments in governance research are standardized tools that generate comparable assessments across organizational units and time periods. A validated instrument must demonstrate content validity (coverage of the construct domain), predictive validity (accuracy in classifying outcomes), and inter-rater reliability (consistency across independent coders or data sources) [4].

2.2 Theoretical Framework

The study integrates two theoretical perspectives. **Prospect theory** provides a behavioral rationale for why firms persist in decoupling: managers may maintain purpose claims because the perceived losses of retracting them (reputational damage, stakeholder backlash) outweigh the perceived gains of honest recalibration, even when practices fail to substantiate the claims. This

asymmetry predicts that decoupling, once established, exhibits inertia—a prediction directly testable through longitudinal trajectory analysis.

Performative purpose alignment theory (PPAT) [2] provides the substantive mechanism: purpose communication is not merely descriptive but creates binding expectations through patterned discursive acts. PPAT specifies three dimensions—semantic alignment with societal agendas, tonal configuration expressing evaluative commitment, and narrative architecture across reporting periods—that jointly determine whether purpose discourse invites or undermines stakeholder inference. For the PPDT, this framework implies that linguistic indicators of purpose statement specificity and consistency should carry predictive weight for subsequent practice trajectories.

2.3 Empirical Review

Jabbouri, Schneckenberg, and Truong [3] conduct a systematic review distinguishing policy–practice from means–ends decoupling, finding that organizations may faithfully implement adopted policies yet still fail to achieve intended outcomes. Their review does not, however, develop measurement instruments or test trajectory dynamics. Powell, Grosvold, and Millington [5] argue that decoupling is often an unintended outcome of attempts to tightly couple rather than a deliberate strategy, proposing conditions under which implementation yields coupled or decoupled outcomes. This insight suggests that trajectory analysis must distinguish between firms that are failing to couple despite effort and those that are not attempting coupling at all.

Fujimura’s [6] study of Japanese higher education organizations applies the means–ends decoupling framework to pedagogical reform, finding that recommended active learning methods had no significant effect on student understanding except in small humanities classes. This finding demonstrates the value of outcome-based decoupling assessment but relies on cross-sectional program evaluation rather than longitudinal trajectory tracking. Okeke [4] develops a multidimensional scale for measuring corporate environmental misrepresentation, validating constructs across misleading claims, selective disclosure, symbolic environmentalism, ESG metric manipulation, and digital greenwashing. While this instrument captures misrepresentation comprehensively, it is designed for cross-sectional assessment and does not model temporal dynamics.

2.4 Research Gap

No validated predictive instrument exists that models purpose–practice decoupling as a longitudinal trajectory using combined disclosure, resource, and operational indicators. Existing tools are either static [4], qualitative [3], or focused on outcomes without tracking the evolution of the claim–practice relationship [6]. The PPDT fills this gap by operationalizing decoupling as a measurable trajectory, validating its predictive accuracy against held-out observations, and specifying the indicator combinations that most reliably forecast persistent divergence.

3. Methodology

3.1 Research Design

The study employs a quantitative, design-based research design combining retrospective panel analysis with prospective simulation. This design is appropriate because the research objective is instrument development and validation rather than causal inference. Retrospective analysis establishes the empirical relationships among indicators and outcomes; prospective simulation tests the instrument's performance on data not used in development, approximating deployment conditions.

3.2 Study Area / Population

The target population comprises publicly listed corporations in the United States and United Kingdom that published standalone sustainability or purpose reports across all three reporting years (2021–2023) and for which financial and operational data were available through standard databases. This population is theoretically appropriate because these firms operate under high disclosure expectations and face institutional pressures to demonstrate purpose–practice alignment.

3.3 Sample Size and Sampling Technique

The final analytic sample comprises 184 firms (552 firm-year observations). Stratified random sampling ensured representation across five sectors: technology, financial services, healthcare, consumer goods, and industrials. Stratification was necessary because decoupling dynamics may vary systematically by sector due to differing regulatory scrutiny and stakeholder salience. Power analysis indicated that 180 firms provide 80% power to detect a medium effect size ($f^2 = 0.15$) in the classification models at $\alpha = 0.05$.

3.4 Data Collection Methods

Three data streams were collected. First, purpose and sustainability report texts were extracted from corporate websites and regulatory filings for each firm-year. Second, financial data (capital expenditure, R&D spending, employee compensation ratios) were obtained from Compustat and Refinitiv. Third, operational indicators (employee turnover, emissions intensity, supplier audit coverage) were collected from sustainability reports and third-party data providers. Where operational indicators were unavailable for more than 15% of firm-years, values were imputed using multiple imputation with chained equations ($M = 20$). For a subset of firms ($n = 23$) with systematically incomplete operational disclosure, simulated values were generated based on sector-year medians with added noise, a procedure necessitated by the instrument's requirement for complete trajectory data.

3.5 Research Instruments

The analytical pipeline was implemented in Python 3.11 using scikit-learn 1.3, pandas 2.0, and spaCy 3.7. Text preprocessing included tokenization, lemmatization, and removal of boilerplate language identified through n-gram frequency analysis. Purpose statement specificity was operationalized using a fine-tuned sentence-level classifier distinguishing generic commitments (“we are committed to sustainability”) from specific, verifiable commitments (“we will reduce Scope 1 emissions by 40% by 2030”). Consistent with Okeke’s [2] emphasis on discursive costliness, specificity scores were weighted by the presence of numerical targets, temporal deadlines, and named accountability mechanisms. Resource–rhetoric ratio divergence was calculated as the absolute difference between the year-over-year change in purpose-related capital allocation (standardized) and the change in purpose discourse volume (standardized).

3.6 Validity and Reliability

Content validity was established through expert review: five scholars with expertise in corporate governance and measurement development rated the initial indicator pool for relevance and coverage (content validity index = 0.89). **Predictive validity** was assessed through held-out validation: the model was trained on 2021–2022 data and tested on 2023 outcomes, with decoupling status determined by an independent composite of stakeholder complaint incidence, regulatory inquiry, and third-party audit findings. **Inter-rater reliability** for the purpose statement specificity coding was assessed by double-coding a random 10% subsample (n = 55 firm-years), yielding Cohen’s $\kappa = 0.84$, indicating strong agreement.

3.7 Data Analysis Techniques

Three models were compared: (a) a logistic regression baseline using only static disclosure-gap indicators (purpose–practice language divergence at a single point in time), (b) a random forest classifier using the full PPDT indicator set, and (c) a gradient boosting classifier (XGBoost) using the same indicators. Performance metrics included accuracy, F1 score, area under the receiver operating characteristic curve (AUC), and temporal lead time (the number of reporting periods before an observable adverse event that the model correctly classified the firm as persistently decoupled). Five-fold cross-validation was used for hyperparameter tuning; final performance was evaluated on the held-out 2023 test set. Feature importance was assessed using SHAP (SHapley Additive exPlanations) values for the gradient boosting model [4].

3.8 Ethical Considerations

All data were obtained from publicly available regulatory filings, corporate sustainability reports, and licensed financial databases. No personally identifiable information or protected health information was accessed. The study used de-identified firm-level data only. Institutional Review Board exemption was granted under 45 CFR 46.104(d)(4) for secondary research on publicly available data.

4. Results

4.1 Data Presentation

Table 1 presents descriptive statistics for the persistently decoupled group (identified by the independent outcome composite) and the tightening-coupling group across key indicators.

Table 1. Key Indicators by Decoupling Status (2021–2023)

Indicator	Persistently Decoupled (n = 71 firms)	Tightening Coupling (n = 113 firms)
Purpose specificity score (mean, SD)	2.14 (0.67)	3.42 (0.81)
Resource–rhetoric divergence (mean, SD)	0.78 (0.24)	0.31 (0.18)
Operational alignment index (mean, SD)	0.42 (0.15)	0.71 (0.14)
Employee turnover ratio (mean, SD)	0.19 (0.07)	0.12 (0.05)
Emissions intensity change (mean, SD)	+2.3% (5.1%)	−8.7% (6.4%)

As shown, persistently decoupled firms exhibit lower purpose specificity, higher resource–rhetoric divergence, and weaker operational alignment across the observation window.

4.2 Analysis of Results

The gradient boosting model achieved the best performance on the held-out 2023 test set, with **89.4% classification accuracy** ($F1 = 0.87$, $AUC = 0.91$). The random forest model achieved 86.2% accuracy, while the static disclosure-gap baseline achieved only 65.8% accuracy, a difference of 23.6 percentage points attributable to the PPDT’s inclusion of resource and operational trajectory indicators. The improvement in F1 from baseline (0.61) to PPDT (0.87) was statistically significant (McNemar’s test, $p < 0.001$).

Feature importance analysis (SHAP values) identified the following top predictors: (1) narrative specificity decay (mean $|\text{SHAP}| = 0.31$), defined as the year-over-year change in the proportion of specific, verifiable commitments; (2) resource–rhetoric ratio divergence (mean $|\text{SHAP}| = 0.27$); (3) operational alignment index trend (mean $|\text{SHAP}| = 0.22$); and (4) temporal coherence of purpose statement architecture (mean $|\text{SHAP}| = 0.14$). The static disclosure-gap indicator contributed only 0.06 to predictive power.

5. Discussion

5.1 Interpretation

The finding that narrative specificity decay is the strongest single predictor of persistent decoupling has direct theoretical implications. It supports Okeke’s [2] performative framework by demonstrating that the linguistic architecture of purpose communication—specifically, the progressive erosion of specific, verifiable commitments—carries predictive information about subsequent practice trajectories. Firms that begin with specific commitments but gradually retreat into generic language are signaling, whether intentionally or not, a decoupling trajectory. This finding answers Research Question 1 by identifying the combination of specificity decay and resource–rhetoric divergence as the most informative indicator pair.

The PPDT’s 89.4% accuracy substantially exceeds the static baseline, directly addressing Research Question 2. The 23.6 percentage point improvement demonstrates that tracking change over time—not merely assessing a snapshot gap—is essential for accurate classification. This extends Bromley and Powell’s [1] conceptual distinction by providing empirical evidence that decoupling trajectories have measurable signatures distinguishable from static states. The finding also aligns with Powell et al.’s [5] argument that decoupling is often an unintended outcome of failed coupling attempts rather than deliberate evasion: the trajectory analysis reveals that some firms begin with specific commitments (indicating genuine coupling intent) but progressively decouple as operational indicators fail to keep pace.

5.2 Implications

Academic implications: The study introduces trajectory-based measurement to the decoupling literature, operationalizing a construct that has remained largely qualitative or cross-sectional. It extends performative purpose theory [2] by demonstrating that discursive patterns have longitudinal predictive validity when combined with material indicators. The finding that specificity decay precedes observable adverse outcomes suggests a temporal ordering that warrants further theoretical elaboration.

Practical implications: Administrators and regulators can deploy the PPDT as a monitoring instrument requiring only publicly available data. Firms exhibiting specificity decay scores

above the 75th percentile and resource–rhetoric divergence above the 60th percentile should be flagged for escalated review, with an expected lead time of one to two reporting periods before adverse events materialize. The instrument’s modular structure allows adaptation to sector-specific indicator sets without full revalidation.

5.3 Limitations

1. The sample is limited to 184 publicly listed firms in developed Anglophone markets, constraining generalizability to other governance contexts and ownership structures.
2. Simulated values for a subset of operational indicators ($n = 23$ firms) may introduce measurement error, although sensitivity analyses excluding these firms yielded comparable accuracy (87.9%).
3. The study assumes stability of historical trajectory patterns; structural shocks or regulatory regime changes could alter the relationship between indicators and outcomes.
4. The outcome composite, while independently constructed, relies on observable adverse events that may not capture all forms of decoupling, particularly those that remain internal to the organization.

5.4 Future Research Directions

- Extension of the PPDT to non-listed firms, state-owned enterprises, and firms in emerging markets where disclosure infrastructures differ substantially.
- Longitudinal qualitative study examining how administrators interpret and respond to PPDT classifications, including potential gaming behaviors.
- Development of sector-specific indicator weights to improve accuracy in industries with distinctive purpose–practice relationships (e.g., extractives, healthcare).
- Integration of the PPDT with real-time digital disclosure monitoring to reduce the temporal lag between decoupling onset and detection.

6. Conclusion

This study developed and validated the Purpose–Practice Decoupling Tracker, demonstrating that decoupling can be measured as a trajectory with 89.4% classification accuracy when disclosure semantics, resource allocation signals, and operational indicators are combined. The instrument’s primary contribution is the operationalization of decoupling as a dynamic process rather than a static state, providing a replicable framework that outperforms static methods by a substantial margin. For administrators, the practical takeaway is that specificity decay and resource–rhetoric divergence are early warning signals requiring attention before reputational or regulatory consequences materialize. As corporate purpose commitments continue to proliferate, the

capacity to track their substantive realization over time becomes not merely an academic exercise but a governance necessity.

References

- [1] Bromley, P., & Powell, W. W. (2012). From smoke and mirrors to walking the talk: Decoupling in the contemporary world. *Academy of Management Annals*, 6(1), 483–530.
- [2] Okeke, A., & Ugbebor, I. (2026). Discursive governance and development goals: A performative theory of corporate purpose in sustainability discourse. *Business Strategy and the Environment*. Advance online publication. <https://doi.org/10.1002/bse.71242>
- [3] Jabbouri, R., Schneckenberg, D., & Truong, Y. (2022). From policy-practice to means-ends decoupling in organizations: A systematic review and paths for future research. *Management International*, 26(1), 123–148.
- [4] Okeke, A. (2026). Green claims and grey areas: A diagnostic tool for assessing corporate environmental misrepresentation. *International Journal of Disclosure and Governance*, 1–17.
- [5] Powell, A., Grosvold, J., & Millington, A. (2015). Trying and failing: An extended conceptual model of decoupling. *Academy of Management Proceedings*, 2015(1), 1–6.
- [6] Fujimura, M. (2017). Analytical perspective on higher education organizations (2): Reform, practice, and outcomes from the decoupling perspective. *Daigaku Ronshu*, 49, 37–52.
- [7] Gionfriddo, G., Testa, F., & Piccaluga, A. (2026). Purpose-washing? The strategic role of rhetoric. *Journal of Business Research*, 189, 115–129.
- [8] Haack, P., & Schoeneborn, D. (2015). Is decoupling becoming decoupled from institutional theory? A commentary on Wijen. *Academy of Management Review*, 40(2), 307–310.
- [9] Jabbouri, R., Truong, Y., & Schneckenberg, D. (2019). Institutional means-ends decoupling work in industrial R&D project implementation. *Industrial Marketing Management*, 80, 45–58.
- [10] Gauch, K., & Theis, J. (2024). Corporate purpose disclosure in the annual report: Substance or cheap talk? An experimental analysis of non-professional investors' perception. *Journal of Accounting and Public Policy*, 45, 107–122.
- [11] Ruiz-Perez, J. (2022). *Purpose-driven organizations: A maturity model for corporate purpose implementation*. Emerald Publishing.
- [12] Björck, F., Leo, U., & Nilsson, A. (2023). Measuring purpose implementation: A multidimensional scale development study. *Corporate Governance*, 23(4), 78–96.
- [13] Etzion, D., & Ferraro, F. (2010). The role of analogy in the institutionalization of sustainability reporting. *Organization Science*, 21(5), 1092–1107.

[14] Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.

[15] Wijen, F. (2014). Means versus ends in opaque institutional fields: Trading off compliance and achievement in sustainability standard adoption. *Academy of Management Review*, 39(3), 302–323.