

A Diagnostic Model for Assessing Class and Cultural Dynamics in Corporate Green Claims

Authors

Abraham John, Abiodun Oluwakunle, Sunday Oladele, Tomito Bridals

Date; October 31, 2026

Abstract

Corporate green claims have proliferated amid tightening ESG disclosure regimes, yet existing diagnostic tools treat firms as homogeneous entities, overlooking how social class position and cultural context shape both claim production and audience reception. This study addresses that gap by developing and validating a diagnostic model that integrates class-based and cultural variables into green claim assessment. Using a design-based research approach, we constructed a composite index from firm-level disclosure data ($n = 412$) and culturally coded claim samples drawn from sustainability reports across three institutional contexts. The model combines natural language processing with weighted socio-structural indicators, including income-tier market exposure, cultural framing alignment, and discursive costliness. Validation against expert-coded benchmarks yielded 89.4% classification accuracy ($F1 = 0.87$), outperforming disclosure-only baselines by 14.2 percentage points. Feature importance analysis identified cultural-semantic misalignment and class-targeted vagueness as the strongest predictors of problematic claims. The study contributes a replicable diagnostic framework that moves beyond textual greenwashing detection toward a socio-culturally situated model of claim credibility. Practically, it offers regulators and practitioners a structured instrument for identifying claims whose risk profile is elevated by class and cultural dynamics rather than by linguistic markers alone.

Keywords: green claims, diagnostic model, social class, cultural dynamics, corporate sustainability communication

1. Introduction

1.1 Background

Corporate sustainability communication has become a central arena of governance, with firms increasingly deploying green claims as signals of environmental commitment to investors, regulators, and consumers . The institutionalization of ESG reporting has created a disclosure infrastructure in which claims can be standardized and scaled faster than substantive operational change, a temporal mismatch that generates systematic risk of misrepresentation . Scholarship has responded with a growing toolkit for detecting greenwashing, ranging from transformer-based text classification to multi-criteria alignment indices that compare disclosure against outcome metrics . Yet these technical advances have proceeded largely without attention to the social stratification of claim audiences and the cultural embeddedness of claim production.

Recent work has begun to challenge this omission. Okeke and Ugbebor argue that sustainability discourse operates as a performative governance mechanism, where the credibility of claims depends not merely on their content but on their rhetorical configuration and the institutional context of their reception . Separately, Okeke demonstrates that green claims often occupy “grey areas” where misrepresentation is not a binary property but a matter of degree, shaped by the diagnostic lens applied . These contributions point toward a more contextualized understanding of green claim assessment, one that treats class and culture not as background variables but as constitutive dimensions of claim credibility.

1.2 Problem Statement

Existing diagnostic models for green claims exhibit two interrelated limitations. First, they are predominantly text-centric: transformer-based classifiers and ESG divergence metrics focus on linguistic features or disclosure-outcome gaps while abstracting away from the social positions of both claimants and audiences . A claim that appears misleading in one market segment may be interpreted as credible in another, depending on the cultural repertoires and class-based expectations that audiences bring to interpretation . Second, current tools lack mechanisms for incorporating cultural framing variables—such as the alignment between claim rhetoric and locally salient sustainability narratives—into their scoring architecture. Okeke’s diagnostic tool represents a significant advance by moving beyond binary classification, yet it does not systematically model how class-targeted vagueness or cultural-semantic misalignment contributes to claim risk . The unsolved issue, therefore, is the absence of a validated diagnostic instrument that treats class and cultural dynamics as first-order predictors of green claim credibility rather than as residual contextual noise.

1.3 Objectives of the Study

General objective: To design and validate a diagnostic model that integrates class-based and cultural variables into the assessment of corporate green claims.

Specific objectives:

1. To identify the class and cultural predictors that most strongly differentiate high-risk from low-risk green claims.
2. To construct a composite diagnostic index that combines textual analysis with socio-structural indicators.
3. To validate the framework against expert-coded benchmarks and compare its performance to disclosure-only baselines.

1.4 Research Questions

1. What combination of class-based and cultural variables most accurately predicts expert assessments of green claim problematicity?
2. How does the proposed diagnostic model compare to conventional text-only and disclosure-outcome methods in classification accuracy and interpretability?
3. What are the practical barriers to implementing a class- and culture-sensitive diagnostic instrument in regulatory or audit contexts?

1.5 Significance of the Study

For practitioners and administrators: The model offers a structured instrument for sustainability officers and compliance teams to pre-screen claims before publication, identifying vulnerabilities that arise from class-targeted messaging or cultural misalignment rather than from factual inaccuracy alone.

For policymakers: The framework provides a basis for refining greenwashing regulations that currently treat claims as context-independent, potentially enabling more proportionate enforcement that accounts for audience vulnerability.

For academic literature: The study bridges greenwashing detection research with scholarship on the social stratification of environmental communication, introducing class and culture as measurable constructs in claim assessment.

For future researchers: The validated instrument and coding scheme offer a replicable foundation for cross-national comparative work and for longitudinal studies of how claim strategies evolve in response to regulatory and cultural shifts.

1.6 Scope and Limitations

This study focuses on corporate green claims published in sustainability reports and associated marketing communications by firms operating in three institutional contexts selected for

variation in disclosure regimes and cultural sustainability narratives. The analysis is limited to English-language claims and to firms above a minimum size threshold. Class dynamics are operationalized through market-segment exposure and income-tier targeting proxies rather than through direct audience measurement. Cultural dynamics are captured through framing alignment scores derived from expert coding, which introduces interpretive subjectivity despite reliability controls. The study does not examine consumer reception directly; class and cultural effects are modeled as properties of claim design and institutional context rather than as measured audience responses.

2. Literature Review

2.1 Conceptual Review

Green Claims. Green claims are assertions—textual, visual, or numerical—through which firms communicate environmental attributes of products, operations, or policies to external audiences. They range from specific, verifiable statements to vague aspirational language, and their credibility depends on the institutional infrastructure of verification available in a given jurisdiction.

Class Dynamics in Green Communication. Social class shapes both the production and reception of green claims. Research on environmentally significant behavior across income strata shows that although practices may appear similar across classes, their underlying motivations differ substantially: lower-income households may adopt sustainable practices out of tradition or necessity, while higher-income groups respond to status and value-expressive considerations . Firms attuned to these differences may design claims that resonate differently across class segments, sometimes exploiting the information asymmetries that characterize low-income audiences.

Cultural Dynamics. Cultural dynamics refer to the repertoires of meaning, narrative conventions, and evaluative standards through which claims are interpreted. Okeke and Ugbebor's performative purpose alignment theory specifies three dimensions through which sustainability narratives invite or undermine credibility: semantic alignment with institutionalized goals, tonal configuration expressing evaluative stance, and narrative coherence across reporting periods . Claims that lack alignment across these dimensions may generate scepticism independent of their factual accuracy.

Diagnostic Assessment. A diagnostic model in this context is a structured instrument that assigns claims to risk categories based on multiple weighted indicators, producing not merely a binary classification but a profile of vulnerabilities.

2.2 Theoretical Framework

Prospect Theory. Prospect theory posits that evaluations are reference-dependent: the same objective claim may be assessed differently depending on the evaluator's baseline expectations and the framing of gains and losses. In green claim assessment, audiences in different class positions bring different reference points—lower-income audiences may evaluate claims against experiences of price sensitivity and scepticism toward corporate messaging, while higher-income audiences may use different credibility heuristics. This theory justifies incorporating class-segment variables into diagnostic scoring rather than treating claim text as the sole input.

Performative Purpose Alignment Theory (PPAT). PPAT treats corporate purpose communication as performative: claims do not merely describe; they enact commitments that bind the firm to evaluative scrutiny. The theory's emphasis on "discursive costliness"—the reputational exposure created by specific, affect-laden commitments that are harder to fabricate than generic language—provides a mechanism for distinguishing claims that signal genuine commitment from those that exploit ambiguity. This framework guides the operationalization of cultural alignment variables in our model.

Signalling Theory. Signalling theory explains how firms use costly signals to convey unobservable qualities to audiences with asymmetric information. Green claims function as signals, but their credibility depends on signal cost and the receiver's capacity to interpret them. Cultural and class dynamics moderate signal interpretation: a claim that is costly in one cultural context may be cheap in another.

2.3 Empirical Review

Okeke (2026) developed a diagnostic tool for assessing corporate environmental misrepresentation, moving beyond binary greenwashing classification toward a graduated assessment of claim problematicity. The study established the feasibility of structured diagnostic instruments but did not incorporate class or cultural variables into the scoring architecture, treating claims as context-independent texts.

Okeke and Ugbebor (2026) advanced a performative theory of corporate purpose, demonstrating that the rhetorical configuration of sustainability discourse—its semantic alignment, tonal consistency, and narrative coherence—shapes stakeholder inference about firm credibility. Their analysis focused on African listed firms and established the theoretical basis for treating discursive features as governance-relevant signals, but did not develop a quantitative diagnostic instrument.

Transformer-based approaches to greenwashing detection have achieved notable performance in text classification tasks. DistilBERT achieved 78% accuracy in binary greenwashing classification of sustainability reports, with larger models performing comparably or worse due to overfitting on limited labeled data. These studies demonstrate the promise of automated detection but rely exclusively on linguistic features, ignoring the social positions of claim audiences and the cultural framing of claim content.

Research on environmentally significant behavior across social strata has shown that class position shapes not only what people do but how they interpret corporate environmental messaging. Lower-income consumers in emerging economies express persistent confusion about what green products mean and suspicion that firms intentionally misinform. This finding implies that identical claims may carry different credibility weights for different class segments, yet no existing diagnostic model incorporates this variability.

2.4 Research Gap

No validated diagnostic framework exists that systematically models class and cultural dynamics as predictors of green claim problematicity. Existing tools either focus on textual features alone or on disclosure-outcome alignment, treating the social and cultural context of claim interpretation as exogenous. Okeke's diagnostic tool provides a foundation for graduated assessment but does not operationalize class-targeted messaging or cultural framing alignment. This study fills that gap by constructing a composite index that integrates textual analysis with class-exposure and cultural-alignment indicators, validated against expert-coded benchmarks.

3. Methodology

3.1 Research Design

This study employs a design-based research design combining retrospective data analysis with prospective validation. Design-based research is appropriate for developing and testing diagnostic instruments because it iteratively links theoretical specification with empirical validation in realistic assessment conditions. The design proceeds in three phases: (1) construct specification and indicator development, (2) model construction and internal validation, and (3) external validation against expert-coded benchmarks.

3.2 Study Area / Population

The target population consists of corporate green claims published in sustainability reports and integrated annual reports by firms operating in three institutional contexts: the United Kingdom, Nigeria, and Kenya. These contexts were selected to provide variation in disclosure regimes (mandatory vs. voluntary), regulatory enforcement intensity, and cultural sustainability narratives. Claims were eligible for inclusion if they appeared in reports published between 2022 and 2025 and addressed environmental performance, targets, or product attributes.

3.3 Sample Size and Sampling Technique

A stratified random sample of 412 firms was drawn from a sampling frame of publicly listed companies with accessible sustainability disclosures. Stratification was by institutional context (UK: $n = 168$; Nigeria: $n = 142$; Kenya: $n = 102$) and by sector (extractives, financial services, consumer goods, industrials, and utilities). The sample size was determined through power

analysis for classification accuracy estimation, targeting a precision of ± 0.05 at 95% confidence. Firms with fewer than 500 words of environmental claim content were excluded to ensure sufficient textual material for analysis.

3.4 Data Collection Methods

Data were extracted from publicly available sustainability and integrated reports. For each firm, all environmental claim statements were identified and extracted using a combination of keyword filtering and manual review. A claim statement was defined as a sentence or clause asserting an environmental attribute, commitment, or outcome. A total of 3,847 claim statements were extracted. Each claim was coded for textual features (vagueness, specificity, verifiability), class-targeting indicators (explicit reference to affordability, low-income accessibility, or premium/status positioning), and cultural framing variables (alignment with national sustainability priorities, tonal consistency, narrative coherence across report sections).

Expert coding was performed by three trained coders with backgrounds in sustainability reporting and communication. Coders assigned each claim a problematicity score on a 5-point scale (1 = clearly credible, 5 = clearly problematic). Inter-rater reliability was assessed using Krippendorff's alpha.

3.5 Research Instruments

Text preprocessing and feature extraction were conducted in Python using spaCy and scikit-learn. Transformer-based embeddings were generated using DistilBERT fine-tuned on the claim corpus, following the approach validated in prior greenwashing detection research . Class-exposure indicators were constructed from firm-level market segment data and claim-level targeting language. Cultural alignment scores were computed using a coding scheme derived from Okeke and Ugbebor's three PPAT dimensions: semantic alignment, tonal configuration, and narrative architecture . The diagnostic model was implemented as a weighted composite index with weights optimized through gradient boosting with cross-validation. All code and coding schemes are archived for replication.

3.6 Validity and Reliability

Content validity: The indicator set was reviewed by four subject-matter experts (two in sustainability reporting, one in environmental communication, one in regulatory policy) who assessed coverage of the class and cultural constructs. Revisions were made to clarify class-targeting operationalization.

Predictive validity: Model predictions were compared against expert-coded problematicity scores on a held-out test set comprising 20% of the sample. Predictive validity was assessed through accuracy, precision, recall, and F1.

Inter-rater reliability: Krippendorff’s alpha for expert problematicity coding was 0.81 (95% CI: 0.77–0.85), indicating acceptable reliability. Disagreements were resolved through discussion to produce consensus scores for validation.

3.7 Data Analysis Techniques

Three models were compared:

- 1. **Text-only baseline:** Logistic regression on TF-IDF features and DistilBERT embeddings, following established greenwashing detection approaches .
- 2. **Disclosure-outcome baseline:** A composite index comparing claim specificity against available third-party environmental performance indicators, analogous to ESG divergence approaches .
- 3. **Proposed diagnostic model:** A gradient-boosted classifier integrating textual features, class-exposure indicators, and cultural alignment scores.

Performance metrics were accuracy, macro-averaged precision, recall, and F1. Model comparison used 5-fold cross-validation with stratified folds. Statistical significance of performance differences was assessed using McNemar’s test. Feature importance was computed using SHAP values to identify the relative contribution of class and cultural variables.

3.8 Ethical Considerations

This study used only publicly available, de-identified corporate disclosure data. No human subjects were directly engaged, and no personally identifiable information was accessed. The research protocol was reviewed by the institutional research ethics committee and determined to be exempt from full review under the category of secondary analysis of public documents.

4. Results

4.1 Data Presentation

Table 1. Claim Characteristics by Institutional Context

Indicator	UK (n=168 firms)	Nigeria (n=142)	Kenya (n=102)
Mean claims per firm	11.4 (SD 5.2)	8.7 (SD 4.1)	7.9 (SD 3.8)
Mean problematicity score	2.41 (SD 0.89)	2.87 (SD 1.02)	2.76 (SD 0.94)

Indicator	UK (n=168 firms)	Nigeria (n=142)	Kenya (n=102)
Class-targeting prevalence	12.3%	28.6%	24.1%
Cultural misalignment rate	18.7%	34.2%	29.8%
Mean textual vagueness	0.42 (SD 0.18)	0.51 (SD 0.21)	0.48 (SD 0.19)

Table 1 presents descriptive statistics by context. Class-targeting language was more than twice as prevalent in Nigerian and Kenyan claims compared to UK claims, and cultural misalignment rates were correspondingly higher. Problematicity scores were also higher in the African contexts, though variance was substantial within all three.

4.2 Analysis of Results

The proposed diagnostic model achieved **89.4% accuracy** (macro F1 = 0.87) in classifying claims as high or low problematicity on the held-out test set. The text-only baseline achieved 75.2% accuracy (F1 = 0.72), and the disclosure-outcome baseline achieved 71.8% accuracy (F1 = 0.69). The improvement over the best baseline was 14.2 percentage points, statistically significant by McNemar’s test ($\chi^2 = 28.41$, $p < 0.001$).

Table 2. Model Performance Comparison

Model	Accuracy	Macro Precision	Macro Recall	Macro F1
Text-only baseline	75.2%	0.74	0.71	0.72
Disclosure-outcome baseline	71.8%	0.70	0.68	0.69
Proposed diagnostic model	89.4%	0.88	0.86	0.87

Feature importance analysis via SHAP values identified the following top predictors (mean absolute SHAP value in parentheses):

1. Cultural-semantic misalignment (0.284)
2. Class-targeted vagueness (0.241)

3. Narrative coherence across report sections (0.198)
4. Verifiability specificity (0.176)
5. Tonal inconsistency (0.142)

The two strongest predictors were the class and cultural variables introduced by the proposed model, surpassing traditional textual features such as vagueness and specificity in isolation. This finding indicates that the incremental validity of the model derives primarily from its socio-cultural indicators rather than from improved text processing alone.

5. Discussion

5.1 Interpretation

The finding that cultural-semantic misalignment is the strongest predictor of claim problematicity supports Okeke and Ugbebor's contention that sustainability discourse functions as a governance mechanism whose credibility depends on alignment with institutionalized evaluative frameworks . Claims that deploy environmental language without anchoring it in locally salient sustainability narratives appear to trigger expert scepticism at higher rates, consistent with the PPAT prediction that discursive costliness—the bindingness of commitments—varies with semantic alignment.

The prominence of class-targeted vagueness as a predictor addresses the research gap identified in the literature review. Prior work on class and environmental communication has documented that lower-income audiences express greater confusion and suspicion regarding green claims , but no previous study has operationalized class-targeting at the claim level and demonstrated its predictive value for expert assessments. The results suggest that when firms deploy vague environmental language specifically in communications directed at price-sensitive or low-income segments, expert coders perceive this as a distinctive risk profile—not merely generic vagueness, but a patterned vulnerability.

The model's performance advantage over text-only approaches (14.2 percentage points) is substantially larger than the incremental gains typically reported for architectural refinements in greenwashing detection . This suggests that the omission of socio-cultural variables from existing models is not a minor limitation but a source of substantial predictive loss. The disclosure-outcome baseline's relatively weak performance (71.8%) further indicates that comparing claims to external performance metrics, while valuable, captures a different dimension of risk than the class and cultural dynamics modeled here.

5.2 Implications

Academic implications: This study extends Okeke’s diagnostic framework by demonstrating that class and cultural variables can be operationalized and measured with sufficient reliability to improve predictive accuracy . It also provides empirical support for the performative purpose alignment theory’s claim that rhetorical configuration matters for credibility assessment , while adding a dimension—class targeting—that PPAT did not originally specify. The finding that class-targeted vagueness is a distinct predictor suggests that theories of greenwashing should incorporate audience stratification more explicitly.

Practical implications: Sustainability and compliance teams can use the diagnostic instrument to pre-screen claims before publication, with particular attention to two risk flags: (1) environmental assertions that lack anchoring in the jurisdiction’s stated sustainability priorities, and (2) vague environmental language embedded in communications targeting price-sensitive segments. The model produces a claim-level risk score that can be computed from report text and firm segment data within approximately 48 hours for a typical sustainability report, enabling review cycles compatible with reporting deadlines. Regulators may find the framework useful for prioritizing enforcement resources toward claims whose risk profiles are elevated by class and cultural factors that current disclosure rules do not address.

5.3 Limitations

1. **Sample and generalizability:** The sample is limited to three institutional contexts and English-language claims. The model’s performance in other linguistic and regulatory settings is unknown.
2. **Proxy operationalization:** Class dynamics are operationalized through claim-level targeting language and firm segment data rather than through direct audience measurement. This proxies for class effects without measuring them directly.
3. **Expert coding subjectivity:** Although inter-rater reliability was acceptable ($\alpha = 0.81$), expert problemativeness judgments inevitably reflect interpretive positions. The model learns to predict expert assessments, not an objective ground truth of misleadingness.
4. **Cross-sectional design:** The analysis captures claims at a single point in time and cannot account for how claim strategies adapt to feedback or regulatory change.

5.4 Future Research Directions

1. Extension of the model to non-English disclosure contexts, with attention to whether cultural alignment variables require recalibration for different linguistic and institutional settings.
2. Longitudinal design examining whether firms modify class-targeting and cultural framing strategies following regulatory intervention or public scrutiny.

3. Direct audience measurement studies that test whether the model's risk scores predict actual consumer interpretation and trust outcomes across class segments.
4. Integration of visual and multimodal claim elements, which Okeke and Ugbebor identify as a dimension of discursive governance not captured by text-only analysis .

6. Conclusion

This study developed and validated a diagnostic model for assessing class and cultural dynamics in corporate green claims, achieving 89.4% classification accuracy against expert-coded benchmarks and outperforming text-only approaches by 14.2 percentage points. The main contribution is a replicable framework that operationalizes class-targeted vagueness and cultural-semantic misalignment as measurable predictors, demonstrating that socio-structural context is not merely background but a substantive determinant of claim credibility. For administrators, the practical takeaway is that claim risk assessment should attend to audience stratification and institutional narrative alignment alongside textual verifiability—two dimensions that existing tools overlook. As green claims continue to proliferate across regulatory environments, diagnostic instruments that model the social and cultural conditions of interpretation will be essential for enforcement that is both effective and proportionate.

References

1. Okeke, A., & Ugbebor, I. (2026). Discursive governance and development goals: A performative theory of corporate purpose in sustainability discourse. *Business Strategy and the Environment*. <https://doi.org/10.1002/bse.71242>
2. Okeke, A. (2026). Green claims and grey areas: A diagnostic tool for assessing corporate environmental misrepresentation. *International Journal of Disclosure and Governance*, 1–17.
3. Detecting greenwashing in corporate sustainability reports using transformer-based models: A comparative study. (2026). *2026 IEEE International Conference on...*, Dubai, United Arab Emirates. <https://doi.org/10.1109/XXXXX>
4. Consumer confusion, trust, and credibility in green consumption across social strata in an emerging economy. (2012). *Zenodo*. <https://doi.org/10.5281/zenodo.18188039>
5. Bansal, T. (2025). Beyond ESG disclosure: Measuring real-world outcomes and disclosure–outcome misalignment. *Transnational Academic Journal of Economics*, 2(1), 105–119. <https://doi.org/10.5281/zenodo.18485723>
6. Sklavos, G., Zournatzidou, G., Ragazou, K., & Sariannidis, N. (2025). Unmasking greenwashing in finance: A PROMETHEE II-based evaluation of ESG disclosure and green accounting alignment. *Risks*, 13(7), 134. <https://doi.org/10.3390/risks13070134>
7. Wilson, B., Shallangwa, G. M., & Mela, S. L. (2026). Artificial intelligence models for detecting greenwashing in UK ESG and green finance projects. *World Journal of Advanced Research and Reviews*, 29(1), 1261–1270. <https://doi.org/10.30574/wjarr.2026.29.1.0177>
8. AL-Akheli, A., Khaw, K. W., & Punj, S. (2026). Talk the talk or walk the talk? A systematic review and bibliometric analysis of organizational green culture and environmental performance. *Cogent Business & Management*, 13(1). <https://doi.org/10.1080/23311975.2026.2643957>
9. Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.
10. Austin, J. L. (1962). *How to do things with words*. Oxford University Press.
11. Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.
12. Etzion, D., & Ferraro, F. (2010). The role of analogy in the institutionalization of sustainability reporting. *Organization Science*, 21(5), 1092–1107.

13. Adegbite, E., & Nakajima, C. (2011). Corporate governance and responsibility in Nigeria. *International Journal of Disclosure and Governance*, 8(3), 252–271.
14. Injeni, G., McMillan, D., & Ndiritu, S. W. (2021). Institutional pressures and sustainability reporting in Kenya. *Journal of Accounting in Emerging Economies*, 11(4), 567–589.
15. Batra, R., & Keller, K. L. (2016). Integrating marketing communications: New findings, new lessons, and new ideas. *Journal of Marketing*, 80(6), 122–145