CAP CONTEMPORARY ACCOUNTING PERSPECTIVES
AN INTERNATIONAL PEER-REVIEWED JOURNAL
REVIEW ARTICLE

VOLUME: 1 • ISSUE: 1 • YEAR • 2026
DOI: https://doi.org/10.68656/865h3273
Pages | 86–105
Submitted | 11 August 2026
Accepted | 18 September 2026
Published | 1 October 2026

Beyond Detection: Explaining Greenwashing Risk through a Theory-Grounded Disclosure Quality Index

Mehmet Günlük[1], Murat Özcan[2], Hakan Bilen[3]

[1] Assoc. Prof. Dr., Milas Vocational School, Muğla Sıtkı Koçman University, Türkiye • ORCID: 0000-0001-9665-7557 • mehmetgunluk@mu.edu.tr
[2] Assoc. Prof. Dr., Faculty of Economics and Administrative Sciences, Bolu Abant İzzet Baysal University, Türkiye • ORCID: 0000-0001-9106-4146 • mozcan@ibu.edu.tr
[3] PhD, Independent Researcher, Türkiye • ORCID: 0000-0002-9482-9519 • bilenhakan14@gmail.com

Suggested Citation: Günlük, M., Özcan, M., & Bilen, H. (2026). Beyond detection: Explaining greenwashing risk through a theory-grounded disclosure quality index. *Contemporary Accounting Perspectives*, 1(1), 86–105. https://doi.org/10.68656/865h3273

Abstract

Purpose: *To propose a conceptual framework for a four-dimensional Disclosure Quality Index capable of distinguishing genuine sustainability disclosure from greenwashing, addressing the gap between decoupling-based quantitative scores and content-analysis-based general disclosure quality indices.*

Methodology: *This is a conceptual/theoretical paper. The proposed index is developed through a synthesis of legitimacy theory, signaling theory, and impression management theory, translated into four measurable dimensions—Volume-Substance Gap, Verifiability and Assurance Level, Linguistic Uncertainty and Rhetorical Strategy (Rhetorical Vagueness), and Temporal Consistency— and operationalized through concrete sub-indicators drawn from established accounting text-analysis methodologies (e.g., FOG Index readability, Loughran-McDonald word lists).*

Findings: *As a conceptual contribution, the paper does not present empirical findings; it offers a theoretically grounded measurement architecture along with a research agenda for future construct- and predictive-validity testing.*

Scope and Limitations: *The theoretical scope is limited to three theories among several plausible alternatives; the relative weighting of the four dimensions is left open for future empirical or expert-panel determination; and the framework draws primarily on regulatory and reporting literature from developed markets.*

Originality / Value: *The paper bridges two previously disconnected streams of the greenwashing measurement literature by providing a theoretically justified, codable index architecture that offers a direct methodological foundation for future empirical validation studies.*

Keywords: Greenwashing; Disclosure Quality Index; Legitimacy Theory; Signaling Theory; Sustainability Reporting

JEL Codes: M41; M14; Q56

1. Introduction

Over the past two decades, sustainability reporting has evolved from a voluntary extension of corporate communications into a central reporting domain that directly influences investor decisions and regulatory oversight (Christensen et al., 2021; Lin et al., 2024). The European Union's Corporate Sustainability Reporting Directive (CSRD) (European Union, 2022) and the International Sustainability Standards Board's (ISSB) IFRS S1/S2 standards (IFRS Foundation, 2023a, 2023b) are compelling companies to disclose their environmental, social, and governance (ESG) performance in an increasingly detailed manner; accordingly, the literature highlights the importance of comprehensive and rigorous scientific evidence aligned with these regulatory objectives (Dinh et al., 2023). However, the fact that disclosure requirements have become mandatory does not automatically lead to an improvement in the quality of disclosures in reports. This mismatch between corporate discourse and actual practice is explained by the concept of "decoupling" in organizational theory (Meyer & Rowan, 1977). According to this concept, organizations can formally conform to the legitimacy expectations of the institutional environment while keeping their actual operational structures largely independent of these expectations. This tendency toward decoupling is also a current and expanding area of research specifically within sustainability reporting; indeed, a systematic review of 74 articles from 44 journals shows that research in this field has increased significantly since 2018 and that the EU's regulatory interventions have directly shaped these decoupling dynamics (Cepêda et al., 2025). The classical concept has been adapted to contemporary organizational contexts through two distinct forms: "means-ends decoupling," which refers to the uncertainty regarding whether the official policies and procedures adopted by organizations actually serve their stated objectives, and "policy-practice decoupling," which refers to the failure of accepted official policies to be reflected in daily operational practices (Bromley & Powell, 2012). This distinction is particularly important in the context of ESG reporting. While regulations such as the CSRD and IFRS S1/S2 compel companies to adopt more comprehensive formal disclosure policies (formal compliance), this does not necessarily mean that these policies actually improve sustainability performance or that the reported information accurately reflects the company's actual operational practices. Indeed, firms under increasing reporting pressure can conceal their actual performance by selectively disclosing relatively harmless impacts to create the appearance of transparency—suggesting that regulatory pressure alone does not eliminate the decoupling between means and ends (Marquis et al., 2016). Similarly, mandatory ESG reporting under the EU's Directive 2014/95 has been found to reduce ESG decoupling among EU-listed firms. The mitigating effect is less pronounced for firms whose ESG information is independently audited (Aboud et al., 2024).

This tendency toward divergence, referred to as "greenwashing" in the sustainability literature, has been theorized as a firm's strategic practice of highlighting positive environmental information while concealing negative information (Delmas & Burbano, 2011; Lyon & Maxwell, 2011). The cost of greenwashing is not merely reputational (Taddeo et al., 2026). Misleading sustainability disclosures can lead to mispricing in capital markets, undermine investment decisions based on ESG ratings, and make it difficult to distinguish the signals from companies undergoing genuine sustainable transformation amid the confusion created by imitators (Berg et al., 2022; Lin et al., 2023; Marquis et al., 2016). For this reason, the ability to systematically and reliably detect greenwashing has become an increasingly important issue for both academia and regulatory

authorities; indeed, a marked increase in the number of systematic reviews on the subject has been observed in recent years (Huang et al., 2025; Lublóy et al., 2025).

In the literature, the identification of greenwashing is evaluated within the framework of two distinct approaches. The first relies on empirical "greenwashing scores" that draw on secondary data sources (e.g., Bloomberg ESG disclosure scores, Refinitiv/Asset4 performance ratings) and quantitatively calculate the discrepancy between the level of disclosure and actual performance. A prominent example of this approach measures the discrepancy between disclosure and performance using a standardized metric (Yu et al., 2020). However, a review covering 111 studies shows that such measurements—including the peer-relative approach described above—have largely been developed independently of one another, are specific to their data sources, and lack standardization; furthermore, the theoretical rationale for selecting specific indicators is often not clearly articulated (Lublóy et al., 2025). Definitions of greenwashing and measurement approaches remain fragmented and have not been integrated under a common theoretical framework (Huang et al., 2025). The second approach consists of descriptive scales and indices that generally assess the quality of sustainability reporting through content analysis based on GRI guidelines (Clarkson et al., 2008). While these scales provide valuable insights into reporting quality, they are not specific to greenwashing; a report that is formally compliant with standards but misleading in content may still score highly on these scales (Boiral, 2013; Michelon et al., 2015). In other words, while existing quality scales are successful at measuring "how much is disclosed," they fall short at distinguishing "how verifiable, concrete, and consistent the disclosed information is."

Based on this, this study proposes a conceptual framework for a Disclosure Quality Index aimed at filling the gap in these two approaches to greenwashing detection. Unlike existing data-driven detection models, the proposed index is built on a four-dimensional structure derived from a synthesis of three established organizational and communication theories. Of these theories, Legitimacy Theory explains how firms' pursuit of social approval shapes their disclosure behavior, while Signaling Theory distinguishes genuine and verifiable information from "cheap talk"—which is cost-free and lacks distinctive power. Impression Management Theory highlights the linguistic and rhetorical strategies managers employ in narrative explanations to create a positive impression among stakeholders (Connelly et al., 2011; Deegan, 2002; Lyon & Maxwell, 2011; Merkl-Davies & Brennan, 2007, 2011). The combined use of these three theories aims to ensure that the proposed index is not only statistically discriminant but also an explainable tool that clarifies why it is discriminant.

This study makes three distinct contributions to the literature. First, it provides the greenwashing detection literature with a solid theoretical foundation that most existing empirical models have largely neglected (Lublóy et al., 2025; Huang et al., 2025). Second, by clarifying the conceptual distinction between sustainability reporting quality scales and the specific detection needs of greenwashing, it provides a roadmap for future scale development efforts. Third, the proposed four-dimensional framework—Volume-Substance Gap, Verifiability and Assurance Level, Linguistic Uncertainty and Rhetorical Strategy, and temporal consistency—is based on concrete and codable indicators that can be empirically tested in the future. As such, the study goes beyond a purely conceptual contribution to provide a direct methodological foundation for subsequent empirical research.

The remainder of the study is structured as follows. Section 2 reviews the existing literature on greenwashing detection and the quality of sustainability reporting, thereby highlighting the aforementioned gap in greater detail. Section 3 discusses the three theories underlying the proposed index and how each justifies a specific dimension of the index. Section 4 details the four-dimensional index structure and its sub-indicators. Section 5 discusses how the index can be applied in practice and which data sources can be used to code it. The sixth and final chapter presents the study's limitations and recommendations for future empirical validation studies.

2. Literature Review

The literature on measuring greenwashing can be grouped into three distinct approaches that have developed largely independently of one another. This section examines each approach separately, discusses their strengths and weaknesses, and ultimately explains why these three approaches have not yet converged into a complementary, integrated framework.

2.1. Quantitative Measures Based on Decoupling

Among the pioneering studies in the literature, the decoupling-scores approach quantitatively calculates the discrepancy between a firm's level of disclosure and its actual environmental/ESG performance (Walker & Wan, 2012; Tashman et al., 2019). Over time, this approach has become one of the most frequently cited in the literature. A prominent example of this approach is a greenwashing score that treats the positive difference between a firm's Bloomberg ESG disclosure level and its ESG performance as measured by independent rating agencies as an indicator of greenwashing—that is, the firm is defined as an "actor that does less than it says" (Yu et al., 2020). A systematic review covering 111 studies demonstrates that while such divergence metrics are frequently used in the literature, there is a significant lack of consensus regarding which disclosure and performance data sources (such as Bloomberg and Refinitiv/Asset4—formerly known as Thomson Reuters ASSET4)—should be used, how data should be standardized, and how industry effects should be controlled (Lublóy et al., 2025). The main advantage of this approach is that it is easily applicable to large-scale, multi-country samples and is reproducible (Tashman et al., 2019). Conversely, the approach's main weakness is that the measure remains entirely dependent on the subjective rating methodology of secondary data providers (Berg et al., 2022; Chatterji et al., 2016; Chatterji et al., 2009) and its failure to provide a direct explanation of which disclosure characteristics (language, verifiability, consistency) indicate greenwashing; in other words, this approach measures greenwashing as an outcome but does not conceptualize the content of the disclosure practices that produce it.

2.2. Content Analysis-Based General Explanatory Quality Indices

The second approach, based on content analysis and with a longer history in the accounting literature (Gray et al., 1995; Hackston & Milne, 1996; Wiseman, 1982), aims to create a quality or comprehensiveness index by directly coding the content of companies' sustainability/environmental reports. A pioneering and still widely cited content-analysis index based on the Global Reporting Initiative (GRI) guidelines systematically scores the scope of voluntary disclosures in companies' environmental and social responsibility reports (Clarkson et al., 2008). This methodology continued to be directly applied in subsequent years and continues to be cited as one of the foundational references in the field in recent meta-analyses (Clarkson et al., 2011; Doan & Sassen, 2020). The strength of such indices lies in their direct focus on the content of disclosures; however, their fundamental limitation is that, while they successfully

measure comprehensiveness (how much is disclosed), they lack a mechanism to distinguish the accuracy, verifiability, or misleading nature of the disclosed information. Indeed, a report that is formally compliant with GRI guidelines can score highly on such indices; this is precisely the scenario that greenwashing enables. Even in reports with the highest level of formal GRI compliance (A/A+), approximately 90 percent of significant adverse events were not reported (Boiral, 2013). Similarly, firms adopting CSR reporting practices such as stand-alone reports, assurance, and reporting guidance do not, on average, provide higher-quality information. Among firms that disclose performance-related information, there is limited evidence that GRI followers are more likely to provide balanced, comparable, and precise disclosures (Michelon et al., 2015).

2.3. Disclosure as a Legitimacy Tool

The third approach focuses on theoretically explaining why disclosure behavior occurs, rather than developing a direct measurement tool (Patten, 1992; O'Donovan, 2002). One of the pioneering studies in this approach, an analysis of the aftermath of the Alaska oil spill found that even firms in the same industry—though not directly responsible for the incident—increased their level of disclosure, thereby revealing that disclosure is a strategic response to a collective legitimacy threat rather than a response to individual firm performance (Patten, 1992). Legitimacy theory was further applied to examine which disclosure strategies firms prefer to close their legitimacy gap (O'Donovan, 2002). Firms with poor environmental performance may allocate more disclosure space to environmental issues in order to restore their legitimacy, a finding interpreted within the framework of legitimacy theory (Cho & Patten, 2007). While this approach is effective in explaining the motivational dynamics underlying greenwashing, it does not provide a directly applicable, codable measurement tool for measuring or identifying greenwashing; the translation of theoretical insights into a concrete index remains largely absent from the literature.

2.4. An Unsynthesized Literature: The Nature of the Gap

An evaluation of these three approaches together clearly reveals the complementary relationship between them and why they do not integrate in their current forms. Discrimination-based measures quantify the extent of the difference but do not explain why or how this difference arises. Content analysis indices, on the other hand, assess the scope of the explanation but do not distinguish its verifiability or linguistic quality. The legitimacy literature, meanwhile, explains motivation but does not translate this into a measurable tool. A review of 180 articles published in 58 journals between 1997 and 2023 also confirms this divergence; definitions of greenwashing, measurement approaches, and theoretical foundations remain fragmented and have not been integrated into a common framework (Huang et al., 2025). A review of empirical greenwashing studies from 1990 to 2022 reaffirms the fragmented nature of these three approaches and shows that researchers remain uncertain about how to operationalize greenwashing (Bernini et al., 2024).

The proposed four-dimensional Disclosure Quality Index aims to bring these three approaches together: the "Volume-Substance Gap" logic drawn from the disaggregation literature, the systematic coding approach drawn from the content analysis literature, and the theoretical rationale derived from legitimacy, signaling, and impression management theories are synthesized into a single conceptual framework. Such an integration effort is not entirely unprecedented in the literature: A five-pillar framework similarly combines quantitative ESG performance data with text analysis of companies' sustainability communications to measure

greenwashing; however, its measurement structure is not grounded in the three theoretical foundations used here (Dorfleitner & Utz, 2024). A review covering 97 articles demonstrates that the greenwashing literature remains fragmented across macro-level theories (legitimacy, organizational, stakeholder), and that the field requires a multi-level, multi-theoretical framework (Forliano et al., 2025). This synthesis is discussed in detail in the following section (Section 3).

3. Theoretical Framework

This section discusses in detail how and why each of the theories—legitimacy theory, signaling theory, and impression management theory—underpins a specific dimension of the proposed index. The aim is to move beyond a list of indicators whose dimensions have been arbitrarily assembled and to establish an integrated structure in which each dimension is grounded in a solid theoretical foundation.

3.1. Legitimacy Theory and the "Volume-Substance Gap" Dimension

Organizational legitimacy can be classified into pragmatic, moral, and cognitive forms, and organizations may use strategic behavior to gain, maintain, or restore it (Suchman, 1995). Within the framework of legitimacy theory adapted to accounting literature, social and environmental disclosures function as a strategic tool to demonstrate to stakeholders that the firm appears to be in alignment with societal expectations (Deegan, 2002). This theoretical proposition points to a directly observable behavioral pattern: a firm seeking legitimacy may find it easier and less costly to increase the volume of its disclosures than to enhance their substantive content. This pattern aligns with the central proposition of the decoupling literature (Bromley & Powell, 2012; Meyer & Rowan, 1977). Therefore, the first dimension of the proposed index—the Volume-Substance Gap—is derived from legitimacy theory: the greater the disparity between the volume of the disclosure and the proportion of concrete, measurable, and verifiable content, the more strongly it indicates that the disclosure is approaching a symbolic strategy aimed at gaining legitimacy.

3.2. Signaling Theory and the "Verifiability and Assurance Level" Dimension

Under conditions of information asymmetry, high-quality actors must send costly and difficult-to-imitate signals in order to distinguish themselves from low-quality actors (Connelly et al., 2011). A game-theoretic model specific to greenwashing applies this logic to environmental disclosures (Lyon & Maxwell, 2011): since disclosures not subject to the threat of independent audit or assurance are virtually cost-free for firms, they can be made with ease even if they do not reflect actual performance—that is, they possess the nature of "cheap talk." In contrast, statements subject to third-party assurance—which are independently verifiable—carry the risk of false claims being exposed and leading to reputational damage; consequently, they are considered more reliable signals that are more strongly linked to actual performance; indeed, voluntarily disclosed environmental statements subject to assurance are of significantly higher quality than those without assurance (Moroney et al., 2012). This theoretical distinction directly justifies the index's second dimension—Verifiability and Assurance Level: whether a statement is subject to third-party assurance, the scope of that assurance (limited/reasonable), and the standard on which it is based (ISAE 3000, AA1000AS) provide a direct way to distinguish whether that statement is a "costly signal" or "cheap talk."

3.3. The Impression Management Theory and the "Linguistic Uncertainty and Rhetorical Strategy (Rhetorical Vagueness)" Dimension

Accounting narratives may involve linguistic and rhetorical strategies that create a positive impression among stakeholders (Merkl-Davies & Brennan, 2007, 2011). These strategies include the use of vague or ambiguous adjectives, conveying information in language that is difficult to interpret, and selective emphasis (highlighting positive outcomes while downplaying negative ones). This claim about readability is supported by evidence from the accounting literature. Li (2008) shows that firms with lower earnings tend to file annual reports that are harder to read, with higher Fog indices and greater length, whereas more readable annual reports are associated with more persistent positive earnings. Similarly, financial-text-specific word lists distinguish categories such as uncertainty, negativity, and caution and reduce systematic misclassification in financial and corporate texts; these tools provide a concrete method for empirically coding this dimension of the proposed index (Loughran & McDonald, 2011). This theoretical proposition directly justifies the index's third dimension—Linguistic Uncertainty and Rhetorical Strategy: the ratio of vague qualitative expressions ("environmentally friendly", "green", "sustainable future") to concrete, measurable commitments that include dates and units can serve as a linguistic indicator of whether a statement stems from an intent to inform or from an intent to manage perceptions.

3.4. The Temporal Dimension of the Dimensions: The Intersection of Legitimacy Restoration and Signal Reliability

The fourth dimension of the index—temporal consistency—is derived not from a single theory but from the interaction of legitimacy and signaling theories over time. Legitimacy must not only be acquired but also maintained and enhanced when threatened (Suchman, 1995); strategies for this enhancement often manifest as the continuous reframing of past commitments (postponing or redefining goals). From the perspective of signaling theory, the reliability of a signal depends not only on its current cost but also on its consistency and frequency over time (Connelly et al., 2011) —constantly postponed or reframed goals erode the signal's discriminatory power. For this reason, a firm's tracking of target-achievement performance across years and the frequency with which targets are reframed are included in the index as indicators of both legitimacy restoration strategies and signal reliability over time.

3.5. Theoretical Integration: Dimension-Theory Mapping

The table below (Table 1) summarizes how the three theories align with the index's four dimensions. This alignment provides a direct foundation for breaking down each dimension into concrete, codable sub-indicators in the following section (Section 4).

Table 1
Theoretical Foundations of the Disclosure Quality Index Dimensions

Dimension	Core Theory	Theoretical Mechanism
1. Volume-Substance Gap	Legitimacy Theory (Suchman, 1995; Deegan, 2002)	The pursuit of symbolic alignment makes increasing volume easier/cheaper than increasing content

Dimension	Core Theory	Theoretical Mechanism
2. Verifiability and Assurance Level	Signaling Theory (Connelly et al., 2011; Lyon & Maxwell, 2011)	The distinction between costly/verifiable signals and "cheap talk"
3. Linguistic Uncertainty and Rhetorical Strategy (Rhetorical Vagueness)	Impression Management Theory (Merkl-Davies & Brennan, 2007, 2011)	Creating a positive impression through ambiguous language and selective framing
4. Temporal Consistency	Legitimacy Repair + Signal Reliability (Suchman, 1995; Connelly et al., 2011)	Postponing/reframing goals over time; can be interpreted as both a legitimacy repair strategy and signal erosion

Note. The table summarizes how each of the four proposed index dimensions is theoretically grounded (see Section 3 for a detailed discussion).

4. Proposed Index Structure

This section provides a detailed examination of the four dimensions whose theoretical foundations were established in Section 3, by breaking each down into concrete and codable sub-indicators. The aim here is to transform the proposed index from a purely conceptual framework into a concrete measurement tool that can be empirically applied in the future through content analysis. Three to four sub-indicators are proposed for each dimension; the decision regarding the final weighting scheme—whether to adopt equal weighting or an expert-based approach such as the Analytic Hierarchy Process—has been deliberately left open and positioned as a methodological choice to be deferred to the empirical validation phase. Throughout Section 4, i denotes the firm and t denotes the reporting year.

4.1. Dimension 1: Volume-Substance Gap

1a. Target Density (TD): This indicator captures the concentration of concrete targets in the sustainability report. Let W_{it} denote the total number of words in the report and $N_{target,it}$ the number of target statements that simultaneously contain a numerical value, a time reference, and a measurable unit. Target Density is the number of concrete targets per thousand words:

$$TD_{it} = \frac{N_{target,it}}{W_{it}} \times 1{,}000, \qquad W_{it} > 0$$

If $W_{it} = 0$ (i.e., no usable text could be extracted), TD_{it} is coded as NA. If $W_{it} > 0$ and $N_{target,it} = 0$, $TD_{it} = 0$ which is a valid (lowest-quality) observation. The inverse measure, the Volume-to-Target Ratio $VTR_{it} = \frac{W_{it}}{N_{target,it}}$ is undefined when the report contains no concrete targets; it is therefore reported only descriptively and is not included in the composite score, since TD and VTR represent the same relationship in opposite directions.

1b. Specificity Share (SS): This indicator measures the share of specific, measurable statements among specific and general statements (e.g., "We will reduce Scope 1 emissions by 30% by 2030 relative to 2020" versus "We are committed to sustainability"). Let $N_{specific,it}$ and $N_{general,it}$ denote the respective counts:

$$SS_{it} = \frac{N_{specific,it}}{N_{specific,it} + N_{general,it}}, \qquad N_{specific,it} + N_{general,it} > 0$$

SS ranges from 0 to 1; values closer to 1 indicate greater specificity. If the report contains neither type of statement SS_{it} is coded as NA.

1c. Visual-Evidence Balance (VEB): Sustainability reports frequently rely on photographs, icons, slogans, and decorative infographics that occupy considerable space without conveying verifiable performance information. VEB measures the share of non-evidentiary visual content within the combined area of non-evidentiary visual elements and evidence-based quantitative displays. Let $A_{visual,it}$ denote the page area occupied by decorative or non-evidentiary visual elements, and $A_{data,it}$ the page area occupied by data tables, charts, and other quantitative displays that report identifiable units, reporting periods, scope, and sources. Charts or infographics that contain verifiable quantitative information are coded as A_{data} not A_{visual}. Areas may be measured in square centimeters or as a percentage of total page area:

$$VEB_{it} = \frac{A_{visual,it}}{A_{visual,it} + A_{data,it}}, \qquad A_{visual,it} + A_{data,it} > 0$$

VEB ranges from 0 to 1; values above 0.5 indicate that decorative visuals occupy more space than evidence-based data displays. Using a share rather than a simple ratio keeps the indicator bounded and avoids an undefined value when a report contains no data tables. If the report contains neither type of element, VEB_{it} is coded as NA. A high VEB may be consistent with symbolic communication, but on its own it is not evidence of greenwashing.

Dimension 1 sub-index. The three indicators are transformed as follows. Here, Nst denotes the prespecified normalized value of each raw indicator, scaled to the [0, 1] interval:

$$Q_{TD,it} = N^{st}(TD_{it}), \quad Q_{SS,it} = N^{st}(SS_{it}), \quad Q_{VEB,it} = 1 - N^{st}(VEB_{it})$$

The Dimension 1 score is the mean of the valid transformed indicators:

$$D_{1,it} = \frac{1}{|V_{1,it}|} \sum_{k \in V_{1,it}} Q_{k,it}\,, \qquad |V_{1,it}| \geq 2$$

Where $V_{1,it}$ is the set of non-missing indicators among Q_{TD}, Q_{SS}, and Q_{VEB}. If fewer than two indicators are available in the valid-indicator set, $D_{1,it}$ is coded as NA. Higher values indicate a smaller gap between disclosure volume and substantive content. Here, k indexes the transformed indicators $Q_{TD,it}$, $Q_{SS,it}$ and $Q_{VEB,it}$ and $|V_{1,it}|$ denotes the number of indicators in $|V_{1,it}|$.

4.2. Dimension 2: Verifiability and Assurance Level

2a. Assurance Presence (AP): Indicates whether the report is subject to third-party assurance (yes/no).

AP ∈ {0, 1}; $AP_{it} = 1$ if the report (or a clearly identified part of it) is subject to independent third-party assurance; otherwise $AP_{it} = 0$.

2b. Scope of Assurance (SA): This indicator records whether the assurance is limited or reasonable. The applicable assurance standard (e.g., ISAE 3000, AA1000AS, or a national auditing standard) is recorded separately.

SA ∈ {0, 1, 2}; $SA_{it} = 0$ for no assurance; $SA_{it} = 1$ for limited assurance, and $SA_{it} = 2$ for reasonable assurance. The assurance standard (ISAE 3000, ISAE 3410, AA1000AS, ISSA 5000, or national standards) is coded as a separate categorical variable and is not included in the score.

2c. Nature of the Assurance Provider (NAP): Indicates whether the entity providing the assurance is an independent audit/assurance firm or a consulting firm affiliated with the company.

NAP ∈ {0, 1}; $NAP_{it} = 1$ if the provider is an independent audit or assurance firm and $NAP_{it} = 0$ if it is a consulting firm affiliated with the company or an internal source.

2d. Data Source Transparency (DST): This indicates whether the calculation methodology for the reported performance data (e.g., emissions calculation protocol) is clearly specified in the report.

DST ∈ {0, 1}; $DST_{it} = 1$ if the calculation methodology or protocol for the reported performance data (e.g., the GHG Protocol for emissions) is explicitly specified; otherwise $DST_{it} = 0$.

Avoiding multiple penalization. In a simple additive formula such as $\frac{\left(AP+\frac{SA}{2}+NAP+DST\right)}{4}$, a report without assurance would be penalized three times, because AP, SA, and NAP would all equal zero. To avoid this, SA and NAP are defined only for assured reports and are combined hierarchically with AP into a single Assurance Block (AB):

$$AB_{it} = AP_{it} \times \left[0.5 \times \frac{SA_{it}}{2} + 0.5 \times NAP_{it} \right]$$

When $AP_{it} = 0$, $AB_{it} = 0$ regardless of SA and NAP, so the absence of assurance is counted only once. When $AP_{it} = 1$, AB_{it} ranges from 0.25 (limited assurance by an affiliated provider) to 1 (reasonable assurance by an independent provider). If $AP_{it} = 1$ but SA or NAP cannot be determined from the assurance statement, AB_{it} is coded as NA.

Dimension 2 sub-index. Within the dimension, the assurance block and data source transparency are combined with equal weights:

$$D_{2,it} = \frac{AB_{it} + DST_{it}}{2}$$

$D_{2,it}$ ranges from 0 to 1. The equal within-dimension weighting is a transparent baseline, not a theoretical claim; alternative within-dimension weights (e.g., a greater weight on the assurance block) should be examined in the sensitivity analyses described in Section 6.2. If either component is NA, $D_{2,it}$ is coded as NA.

4.3. Dimension 3: Linguistic Uncertainty and Rhetorical Strategy (Rhetorical Vagueness)

In operationalizing this dimension, two established tools from the accounting text analysis literature can be utilized: the FOG Index-based readability measure and uncertainty and negativity word lists developed specifically for financial texts (Li, 2008; Loughran & McDonald, 2011). All four indicators in this dimension are negative indicators (higher raw values indicate lower quality). They are normalized using Nst and reversed before aggregation.

3a. Readability Index (FOG): Text complexity, as calculated using the FOG Index or a similar readability measure, serves as an indicator—of whether negative information is conveyed in more

complex language. Firms with lower earnings tend to file harder-to-read and longer annual reports, whereas easier-to-read reports are associated with more persistent positive earnings (Li, 2008).

It is calculated using the formula: FOG=0.4 × [(W / S) + 100 × (K / W)]. Here, W=the total number of words in the text, S=the total number of sentences, and K=the number of words with three or more syllables ("complex" words). As the FOG value increases, the text becomes more difficult to read.

3b. Uncertainty Word Ratio (UWR): Based on the Loughran–McDonald uncertainty word list, this is the ratio of ambiguous expressions in the text to the total number of words (Loughran & McDonald, 2011) and is calculated using the formula:

$$UWR_{it} = \frac{N_{uncertainty,it}}{W_{it}} \times 100, \qquad W_{it} > 0$$

If $W_{it} = 0$, or if a validated dictionary is not available for the language of the report, UWR_{it} is coded as NA. Here, $N_{uncertainty,it}$ denotes the number of uncertainty words identified in the validated dictionary, and W_it denotes the total number of words in the report.

3c. Qualitative Adjective Density (QAD): This is the ratio of unmeasurable qualitative adjectives—such as "green," "environmentally friendly," and "sustainable" —to the total number of such adjectives and their linked supporting evidence. QAD is calculated using the formula:

$$QAD_{it} = \frac{N_{adjective,it}}{N_{adjective,it} + N_{support,it}}, \qquad N_{adjective,it} + N_{support,it} > 0$$

$N_{support,it}$ counts only supporting evidence that is linked to a specific adjective under a pre-specified coding rule (e.g., within the same sentence or paragraph). QAD ranges from 0 to 1; values closer to 1 indicate less substantiated language. If the denominator is zero, QAD_{it} is coded as NA. Here, $N_{adjective,it}$ denotes the number of unmeasurable qualitative adjectives in the report.

3d. Selective Framing (SF): This involves comparing the frequency with which positive performance indicators are highlighted in the report to whether negative indicators receive similar visibility.

Let $G_{positive,it}$ and $G_{negative,it}$ denote composite visibility scores based on pre-specified components such as page position, heading level, typographic emphasis, and the number of words allocated:

$$SF_{it} = \frac{G_{positive,it} - G_{negative,it}}{G_{positive,it} + G_{negative,it}}, \qquad G_{positive,it} + G_{negative,it} > 0$$

SF ranges from −1 to +1; values approaching +1 indicate that positive information is systematically highlighted and negative information is downplayed or omitted. If the denominator is zero, SF_{it} is coded as NA.

Dimension 3 sub-index. All four indicators are normalized and reversed as defined above:

$$Q_{k,it} = 1 - N^{st}(k_{it}), \qquad k \in \{FOG,\ UWR,\ QAD,\ SF\}$$
$$D_{3,it} = \frac{1}{|V_{3,it}|} \sum_{k \in V_{3,it}} Q_{k,it}\,, \qquad |V_{3,it}| \geq 3$$

Where $V_{3,it}$ is the set of non-missing transformed indicators. If fewer than three indicators are available, $D_{3,it}$ is coded as NA. Higher values indicate less ambiguous and less selectively framed language. Here, $N^{st}(k_{it})$ denotes the [0, 1]-normalized raw value for indicator k; k indexes FOG, UWR, QAD, and SF; and $|V_{3,it}|$ denotes the number of non-missing transformed indicators in $V_{3,it}$.

4.4. Dimension 4: Temporal Consistency

The indicators in this dimension require comparison of consecutive reports. A minimum observation window of two consecutive comparable reporting years $(t-1,\ t)$ is required to compute $D_{4,it}$. Three or more consecutive years are needed to examine multi-year patterns.

4a. Target-Achievement Tracking (TAT): Rather than a single yes/no code, TAT is measured as the proportion of prior targets whose achievement status is explicitly reported. Let $N_{prior,it}$ denote the number of quantified targets disclosed in the report for year $t-1$ that are still due or whose deadline has passed, and $N_{reported,it}$ the number of those targets whose status (achieved, on track, not achieved, with figures) is reported in year t:

$$TAT_{it} = \frac{N_{reported,it}}{N_{prior,it}}, \qquad N_{prior,it} > 0$$

If the prior report contained no quantified targets $N_{prior,it} = 0$, TAT_{it} is coded as NA; this is not treated as a failure.

4b. Target Re-framing Frequency (TRF): TRF measures how often targets disappear or are materially changed without explanation. To keep the indicator annual rather than cumulative, it is calculated relative to the targets disclosed in the immediately preceding year. Let $N_{replaced,it}$ denote the number of targets disclosed in year $t-1$ that, in year t, are removed, replaced, re-based, or materially redefined (e.g., a changed base year, scope, or deadline) *without* an explicit explanation and reconciliation:

$$TRF_{it} = \frac{N_{replaced,it}}{N_{prior,it}}, \qquad N_{prior,it} > 0$$

Targets that are achieved and retired, or changed with an explicit and reconciled explanation, are not counted in $N_{replaced,it}$. If $N_{prior,it} = 0$, TRF_{it} is coded as NA. The coding rule identifies an unexplained loss of comparability; it does not infer managerial intent. Multi-year patterns can be examined by averaging TRF_{it} over the observation window.

4c. Methodological Consistency (MC): This refers to whether the performance calculation methodology has changed over the years and whether any methodological changes have been explained retroactively.

MC ∈ {0, 1}; $MC_{it} = 1$ if the performance calculation methodology is unchanged relative to year $t-1$, or if a change is disclosed and prior-year figures are restated or reconciled; $MC_{it} = 0$ if a

change occurs without such explanation. If the methodology cannot be identified in either year, MC_{it} is coded as NA.

Dimension 4 sub-index: TAT and MC are positive indicators, whereas TRF is a negative indicator. If any of the three indicators is NA, $D4_{it}$ is coded as NA.

$D_{4,it} = (TAT_{it} + (1 - TRF_{it}) + MC_{it}) / 3$

4.5. General Structure of the Index

The proposed index (DQI) can be constructed as a weighted sum of the four dimension scores or by reporting them separately by dimension. This structure is mathematically expressed by the following formula:

$$DQI_{it} = \sum_{d=1}^{4} w_d \times D_{d,it}, \qquad \sum_{d=1}^{4} w_d = 1, \quad w_d \geq 0$$

Where DQI_{it} is the disclosure quality score of firm i in year t, $D_{d,it}$ is the quality-oriented score of dimension d (the dimension-score formulas defined above), and w_d is the weight assigned to dimension d. Because all dimension scores range from 0 to 1 and point in the same direction, DQI_{it} also ranges from 0 to 1, and higher values indicate higher disclosure quality. A low DQI value should be interpreted as a signal of elevated disclosure-based greenwashing risk that warrants further examination, not as proof of greenwashing; establishing actual greenwashing requires independent evidence on real performance.

The choice of between-dimension weights is deliberately left open. For transparency and replicability, equal weights $w_1 = w_2 = w_3 = w_4 = 0.25$ are proposed only as a reference specification:

$DQI_{it}^{EW} = 0.25 \times (D_{1,it} + D_{2,it} + D_{3,it} + D_{4,it})$ (EW denotes the equal-weight specification.)

Whether the final weights should remain equal or be determined via an expert panel using the Analytic Hierarchy Process (AHP) or a Delphi study is left to the empirical validation phase (Section 6.2). If any dimension score is NA, the composite DQI is not calculated in the reference specification; the four-dimension scores should then be reported separately, and any re-weighted index must disclose which dimensions were available. Reporting the four-dimension scores alongside the composite is recommended in all cases, since the compensatory structure of the composite aggregation formula allows a weak dimension to be offset by a strong one.

5. Discussion

This section addresses how the Disclosure Quality Index, as specified in Section 4, can be applied empirically: which data sources can be used for coding the index (Section 5.1), how the reliability of the coding process can be ensured (Section 5.2), and the methodological limitations that may be encountered during implementation (Section 5.3) are discussed in turn. This discussion aims to demonstrate that the index is designed not only to be conceptually consistent but also to serve as a tool that can be coded on actual datasets; a broader research agenda regarding empirical validation is addressed in Section 6.2.

5.1. Potential Data Sources

Three main types of primary data sources are envisaged for coding the index:

Independent sustainability reports / integrated reports — Reports prepared in accordance with GRI standards, the ISSB's IFRS S1/S2 framework, the European Sustainability Reporting Standards (ESRS) under the CSRD in the European Union, or national regulatory requirements serve as the primary source for coding both Scope 1 and Scope 3 indicators; Mandatory reports under the CSRD/ESRS may also be used in sample stratification, taking into account the distinction between voluntary and mandatory reporting.

Assurance statements — An independent assurance statement included in the report's appendix or published as a separate document enables the direct coding of Dimension 2; whether the statement exists, its scope (limited/reasonable), the standard on which it is based (ISAE 3000, AA1000AS), and the nature of the assurance provider — that is, whether it is an independent audit or assurance firm or a consulting firm affiliated with the company — can be derived directly from this document.

Multi-year report series — A single year's report is insufficient for coding Dimension 4; a panel structure spanning at least 3–5 years is necessary to identify patterns in the reframing of targets and methodological consistency; keeping the panel too short increases the risk of failing to identify an existing pattern in Dimension 4 indicators.

In addition to these three primary sources, structured ESG databases produced by providers such as Refinitiv/LSEG, Bloomberg, and MSCI can be used as a construct validity check independent of the index's primary coding; however, it must be emphasized that these secondary sources carry their own methodological assumptions and classification criteria; therefore, they cannot replace primary source coding and should be evaluated solely as complementary tools for triangulation.

An approach that evaluates annual report narratives using a multidimensional descriptive profile and a set of criteria can serve as a direct reference point for designing the coding protocol of the proposed index; it demonstrates how quantitative content analysis can be integrated with narrative elements that require subjective judgment (Beattie et al., 2004).

5.2. Coding Process and Reliability

Most of the indicators in Dimensions 1, 2, and 4 can be identified through content analysis coding that is relatively structured and requires a low degree of subjectivity; since these indicators largely require binary (present/absent) or categorical coding, high inter-coder reliability is expected. In contrast, some indicators in Dimension 3—particularly the density of qualitative adjectives in 3c and selective framing in 3d—involve a higher degree of subjective judgment, as the assessment of whether a statement is supported by concrete data or whether an indicator is highlighted depends on the coder's interpretation.

Therefore, in the empirical application of the index, (i) at least two independent coders should be used; (ii) a pilot coding round should be conducted using a detailed codebook prior to coding; (iii) inter-coder reliability should be reported using the Krippendorff alpha coefficient, with reference to the generally accepted threshold of $\alpha \geq 0.80$ ($\alpha \geq 0.667$ as the threshold for limited acceptability) (Hayes & Krippendorff, 2007), and (iv) resolving coding disagreements through arbitration by a third coder. To further reduce subjectivity, it is recommended that the quantitative components of

Dimension 3 (FOG Index and uncertainty-word ratio (UWR)) be calculated using automated text analysis tools (e.g., the Loughran-McDonald dictionary) as much as possible, thereby minimizing the need for human coding.

5.3. Limitations Related to the Scope of Application

There are four foreseeable practical challenges regarding the index's applicability: (i) companies publish reports in different languages; since text analytics tools (FOG Index, Loughran-McDonald dictionary) were developed primarily for English texts, they require language-specific adaptation in multinational samples; (ii) the risk that Dimension 2 will systematically yield low scores and have weak discriminatory power among small and medium-sized enterprises, as these firms often do not obtain independent assurance services; (iii) the lack of multi-year data sets required to identify the reframing of targets (Dimension 4) for newly established firms or those that have recently begun reporting; and (iv) the risk that, should the index gain widespread adoption in the academic literature, firms might learn the coding criteria and adapt their reporting practices to align with the index's metrics—that is, the risk that the measurement itself becomes a target, thereby eroding construct validity. This final risk necessitates that the index be periodically updated over time and that the coding criteria be interpreted not in isolation but in conjunction with changes in firm behavior.

6. Conclusion

6.1. Limitations of the Conceptual Framework

This study, by its very nature, has certain limitations. First, while the three proposed theoretical foundations (legitimacy, signaling, and impression management theories) are comprehensive, they are not exhaustive; other approaches within organizational theory (e.g., stakeholder theory, agency theory, and organizational theory in a broader sense) could have suggested alternative or complementary dimensions for the index; the selection of these three theories is based on their status as the most frequently cited theoretical foundations in the greenwashing literature, but they are not presented as a definitive list.

Second, the fact that the weights to be assigned to the four dimensions (Section 4.5) have not been determined in this study and have been left open constitutes a significant methodological decision point in the index's final empirical application. Another issue closely related to this limitation is that the summation formula defined in Section 4.5 ($DQI_{it} = \Sigma_{d=1}^{4} w_d \times D_{d,it}$) implicitly assumes a compensatory structure: in this structure, a low score in one dimension can be offset by a high score in another dimension. However, given the nature of greenwashing, it may be theoretically more defensible to treat an extremely low score on, for example, the Verifiability and Assurance Level dimension (Dimension 2) as a threshold or disqualifying factor for the index, despite high performance on other dimensions. Which of the compensatory and non-compensatory aggregation rules better reflects the validity of the greenwashing framework is a question that falls outside the scope of this study but must be directly addressed during the empirical validation phase.

Third, the index proposal is largely based on literature derived from regulatory and reporting practices in developed markets (the U.S., the EU); since the prevalence of assurance services, the nature of regulatory pressure, and reporting culture may differ in emerging markets, adapting the index to these contexts requires special attention. In particular, the risk that Dimension 2's

discriminatory power may weaken in markets where independent assurance services are relatively scarce is another issue addressed in Section 5.3.

6.2. Future Empirical Research Agenda

As a natural extension of this conceptual framework, at least four distinct empirical research methodologies can be proposed:

(i) Construct validity testing — The extent to which the proposed four dimensions overlap with or diverge from existing differentiation-based greenwashing scores (e.g., Yu et al., 2020) should be tested; this will clarify whether the proposed index captures a unique structure that cannot be reduced to existing measures, or whether it represents a different operationalization of the same underlying structure.

(ii) Predictive validity test — The extent to which the proposed index predicts a firm's cost of equity capital can be tested against prior evidence on the relationship between disclosure level and the cost of equity capital (Botosan, 1997).

(iii) Weighting and dimension prioritization study — An Analytic Hierarchy Process (AHP) or Delphi study conducted with the participation of an expert panel could ensure that the weights assigned to the four dimensions are grounded in an empirically and theoretically justified basis; this study could also test the preference for the compensated/uncompensated aggregation rule discussed in Section 6.1 against expert opinion.

(iv) Qualitative/mixed-method preparer-side research — Semi-structured interviews with finance/accounting managers, investor relations specialists, and independent assurance providers involved in the sustainability reporting process can help determine to what extent the patterns measured by the index (Volume-Substance Gap, selective framing, etc.) reflect a deliberate strategic choice in internal decision-making processes or a partially unconscious response to corporate pressures; this is a line of research that complements content analysis by directly testing the implicit "intent" assumptions of legitimacy and impression management theories.

6.3. Conclusion

This study has highlighted a long-standing gap in the literature on greenwashing detection: existing measurement approaches are either data-driven and lack a theoretical foundation, or they measure general reporting quality without capturing the Verifiability and Assurance Level and Temporal Consistency dimensions specific to greenwashing. The proposed four-dimensional Disclosure Quality Index—based on a synthesis of legitimacy theory, signaling theory, and impression management theory—offers a theoretically grounded, codable framework to fill this gap.

The study's contribution is not limited to a conceptual proposal; the concrete sub-indicators, measurement metrics, and implementation roadmap detailed in Sections 4 and 5 provide a direct foundation for subsequent empirical research. The limitations discussed in Section 6.1—theoretical scope, weighting uncertainty, and contextual transferability—are positioned as a conscious and honest starting point for this empirical agenda.

From a practical perspective, the proposed index is intended not only for the academic literature but also for regulators (in monitoring the effectiveness of CSRD/ESRS and ISSB standards),

independent assurance providers (as a checklist that can be used to prioritize the scope of assurance regarding coding criteria), and the investor/analyst community (as a complementary layer of verification specific to greenwashing, supplementing existing ESG ratings). As the regulatory importance of sustainability reporting increases, the need for such theoretically sound and empirically applicable measurement tools will continue to grow.

Declarations

Ethics Approval	**This study is conceptual/theoretical in nature and does not involve human or animal subjects or primary data collection; therefore, ethics committee approval was not required.**
Conflict of Interest	The author(s) declare that there is no conflict of interest.
Funding	This research did not receive any specific grant from funding agencies in the public, commercial, or nonprofit sectors.
Author Contributions	Conceptualization: M.G., M.Ö.; Methodology (theoretical framework): M.G., M.Ö.; Literature Review: H.B.; Writing – Original Draft: M.G., M.Ö., H.B.; Writing – Review & Editing: M.G., M.Ö.
Data Availability Statement	Not applicable. This conceptual paper does not use or generate empirical datasets.
Use of AI and AI-Assisted Technologies	The authors used generative artificial intelligence (AI) as a language-support tool to improve the clarity, grammar, and readability of the manuscript. The authors reviewed and edited the AI-generated output as necessary and take full responsibility for the content of the published work.

References

Aboud, A., Saleh, A., & Eliwa, Y. (2024). Does mandating ESG reporting reduce ESG decoupling? Evidence from the European Union's Directive 2014/95. Business Strategy and the Environment, 33(2), 1305–1320. https://doi.org/10.1002/bse.3543

Beattie, V., McInnes, B., & Fearnley, S. (2004). A methodology for analyzing and evaluating narratives in annual reports: A comprehensive descriptive profile and metrics for disclosure-quality attributes. Accounting Forum, 28(3), 205–236. https://doi.org/10.1016/j.accfor.2004.07.001

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. Review of Finance, 26(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Bernini, F., Giuliani, M., & La Rosa, F. (2024). Measuring greenwashing: A systematic methodological literature review. Business Ethics, the Environment & Responsibility, 33(4), 649–667. https://doi.org/10.1111/beer.12631

Boiral, O. (2013). Sustainability reports as simulacra? A counter-account of A and A+ GRI reports. Accounting, Auditing & Accountability Journal, 26(7), 1036–1071. https://doi.org/10.1108/AAAJ-04-2012-00998

Botosan, C. A. (1997). Disclosure level and the cost of equity capital. The Accounting Review, 72(3), 323–349. https://doi.org/10.2308/TAR-9709240185

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. https://doi.org/10.5465/19416520.2012.684462

Cepêda, C., Monteiro, A. P., & Aibar-Guzmán, B. (2025). Decoupling in sustainability reporting: A systematic literature review. Corporate Social Responsibility and Environmental Management, 32(3), 2988–3007. https://doi.org/10.1002/csr.3114

Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors, and strategy researchers. Strategic Management Journal, 37(8), 1597–1614. https://doi.org/10.1002/smj.2407

Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? Journal of Economics & Management Strategy, 18(1), 125–169. https://doi.org/10.1111/j.1530-9134.2009.00210.x

Cho, C. H., & Patten, D. M. (2007). The role of environmental disclosures as tools of legitimacy: A research note. Accounting, Organizations and Society, 32(7–8), 639–647. https://doi.org/10.1016/j.aos.2006.09.009

Christensen, H. B., Hail, L., & Leuz, C. (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. Review of Accounting Studies, 26(3), 1176–1248. https://doi.org/10.1007/s11142-021-09609-5

Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relationship between environmental performance and environmental disclosure: An empirical analysis. Accounting, Organizations and Society, 33(4–5), 303–327. https://doi.org/10.1016/j.aos.2007.05.003

Clarkson, P. M., Overell, M. B., & Chapple, L. (2011). Environmental reporting and its relation to corporate environmental performance. Abacus, 47(1), 27–60. https://doi.org/10.1111/j.1467-6281.2011.00330.x

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. Journal of Management, 37(1), 39–67. https://doi.org/10.1177/0149206310388419

Deegan, C. (2002). Introduction: The legitimizing effect of social and environmental disclosures – A theoretical foundation. Accounting, Auditing & Accountability Journal, 15(3), 282–311. https://doi.org/10.1108/09513570210435852

Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. California Management Review, 54(1), 64–87. https://doi.org/10.1525/cmr.2011.54.1.64

Dinh, T., Husmann, A., & Melloni, G. (2023). Corporate sustainability reporting in Europe: A scoping review. Accounting in Europe, 20(1), 1–29. https://doi.org/10.1080/17449480.2022.2149345

Doan, M. H., & Sassen, R. (2020). The relationship between environmental performance and environmental disclosure: A meta-analysis. Journal of Industrial Ecology, 24, 1140–1157. https://doi.org/10.1111/jiec.13002

Dorfleitner, G., & Utz, S. (2024). Green, green, it's green they say: A conceptual framework for measuring greenwashing at the firm level. Review of Managerial Science, 18, 3463–3486. https://doi.org/10.1007/s11846-023-00718-w

European Union. (2022). Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting (Text with EEA relevance). Official Journal of the European Union, L 322, 15–80. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022L2464

Forliano, C., Battisti, E., de Bernardi, P., & Kliestik, T. (2025). Mapping the greenwashing research landscape: A theoretical and field analysis. Review of Managerial Science, 19, 3407–3456. https://doi.org/10.1007/s11846-025-00856-3

Gray, R., Kouhy, R., & Lavers, S. (1995). Corporate social and environmental reporting: A review of the literature and a longitudinal study of UK disclosure. Accounting, Auditing & Accountability Journal, 8(2), 47–77. https://doi.org/10.1108/09513579510146996

Hackston, D., & Milne, M. J. (1996). Some determinants of social and environmental disclosures in New Zealand companies. Accounting, Auditing & Accountability Journal, 9(1), 77–108. https://doi.org/10.1108/09513579610109987

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. Communication Methods and Measures, 1(1), 77–89. https://doi.org/10.1080/19312450709336664

Huang, Z., Shi, Y., & Jia, M. (2025). Greenwashing: A systematic literature review. Accounting & Finance, 65(1), 857–882. https://doi.org/10.1111/acfi.13352

IFRS Foundation. (2023a). IFRS S1 General requirements for disclosure of sustainability-related financial information. International Sustainability Standards Board. https://www.ifrs.org/issued-standards/ifrs-sustainability-standards-navigator/ifrs-s1-general-requirements/

IFRS Foundation. (2023b). IFRS S2 Climate-related disclosures. International Sustainability Standards Board. https://www.ifrs.org/issued-standards/ifrs-sustainability-standards-navigator/ifrs-s2-climate-related-disclosures/

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics, 45(2–3), 221–247. https://doi.org/10.1016/j.jacceco.2008.02.003

Lin, X., Zhu, H., & Meng, Y. (2023). ESG greenwashing and equity mispricing: Evidence from China. Finance Research Letters, 58, 104606. https://doi.org/10.1016/j.frl.2023.104606

Lin, Y., Shen, R., Wang, J., & Yu, Y. J. (2024). Global evolution of environmental and social disclosure in annual reports. Journal of Accounting Research, 62(5), 1941–1988. https://doi.org/10.1111/1475-679X.12575

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Lublóy, Á., Keresztúri, J. L., & Berlinger, E. (2025). Quantifying firm-level greenwashing: A systematic literature review. Journal of Environmental Management, 373, 123399. https://doi.org/10.1016/j.jenvman.2024.123399

Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate environmental disclosure under threat of audit. Journal of Economics & Management Strategy, 20(1), 3–41. https://doi.org/10.1111/j.1530-9134.2010.00282.x

Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. Organization Science, 27(2), 483–504. https://doi.org/10.1287/orsc.2015.1039

Merkl-Davies, D. M., & Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: Incremental information or impression management? Journal of Accounting Literature, 26, 116–196.

Merkl-Davies, D. M., & Brennan, N. M. (2011). A conceptual framework of impression management: New insights from psychology, sociology, and critical perspectives. Accounting and Business Research, 41(5), 415–437. https://doi.org/10.1080/00014788.2011.574222

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 83(2), 340–363. https://doi.org/10.1086/226550

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. Critical Perspectives on Accounting, 33, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003

Moroney, R., Windsor, C., & Aw, Y. T. (2012). Evidence of assurance enhancing the quality of voluntary environmental disclosures: An empirical analysis. Accounting & Finance, 52(3), 903–939. https://doi.org/10.1111/j.1467-629X.2011.00413.x

O'Donovan, G. (2002). Environmental disclosures in the annual report: Extending the applicability and predictive power of legitimacy theory. Accounting, Auditing & Accountability Journal, 15(3), 344–371. https://doi.org/10.1108/09513570210435870

Patten, D. M. (1992). Intra-industry environmental disclosures in response to the Alaskan oil spill: A note on legitimacy theory. Accounting, Organizations and Society, 17(5), 471–475. https://doi.org/10.1016/0361-3682(92)90042-Q

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of Management Review, 20(3), 571–610. https://doi.org/10.5465/amr.1995.9508080331

Taddeo, S., Regoli, A., Weber, O., & Carè, R. (2026). Green talk, costly walk: The financial cost of greenwashing. Business Strategy and the Environment, 35(6), 8589–8609. https://doi.org/10.1002/bse.70595

Tashman, P., Marano, V., & Kostova, T. (2019). Walking the walk or talking the talk? Corporate social responsibility decoupling in emerging market multinationals. Journal of International Business Studies, 50, 153–171. https://doi.org/10.1057/s41267-018-0171-7

Walker, K., & Wan, F. (2012). The harm of symbolic actions and greenwashing: Corporate actions and communications on environmental performance and their financial implications. Journal of Business Ethics, 109, 227–242. https://doi.org/10.1007/s10551-011-1122-4

Wiseman, J. (1982). An evaluation of environmental disclosures made in corporate annual reports. Accounting, Organizations and Society, 7(1), 53–63. https://doi.org/10.1016/0361-3682(82)90025-3

Yu, E. P., Van Luu, B., & Chen, C. H. (2020). Greenwashing in environmental, social, and governance disclosures. Research in International Business and Finance, 52, 101192. https://doi.org/10.1016/j.ribaf.2020.101192