Accounting, Organizations and Society 115 (2025) 101616

Contents lists available at ScienceDirect

Accounting, Organizations and Society

journal homepage: www.elsevier.com/locate/aos

Is a picture worth a thousand words? Image usage in ESG reports[☆]

Disen Huang Albrecht [a,*], Divya Anantharaman [b,**], Keyi Zhao [c]

[a] Rutgers University, 100 Rockafeller Rd, Piscataway, NJ, 08854, United States
[b] Rutgers University, 1 Washington Pl, Newark, NJ, 07102, United States
[c] Southwestern University of Finance and Economics, 555 Liutai Avenue, 611130, Chengdu City, Sichuan Province, People's Republic of China

ABSTRACT

We collect a novel dataset on images in Environmental, Social, and Governance (ESG) reports using deep learning and artificial intelligence (AI), to test hypotheses arising from experimental literature that visual impression management enhances audience perception. We further combine AI and manual methods to contrast images that lack information content ("generic images" evoking broad ESG themes) against images with information content ("specific images" depicting firm-specific ESG activities). We hypothesize and find that firms (1) operating in socially problematic industries (e.g., "sinful", polluting, or controversial industries), (2) issuing less extensive textual ESG disclosures, and (3) experiencing poor ESG performance tend to use more images, especially generic images, consistent with strategic motivations. We then examine the impact of image usage on stakeholders, observing unduly enhanced perception in retail traders but seldom in institutional investors. We observe that ESG award committees reward specific images and penalize generic images. Our results show that ESG report visual impression management affects stakeholders in the real world, though its impact can be overcome by attention and expertise.

1. Introduction

Glossy company annual reports ubiquitously present images along with words and numbers. Yet, we know little about the impact of these images on market participants. Experimental evidence shows that audiences form more positive impressions of disclosures with images than without images (Bansal & Kistruck, 2006; Cho et al., 2009; Elliott et al., 2017), and small-scale archival studies show that firms strategically use images to enhance audience perception (Hrasky, 2012).

Firms' attempts to "manage impressions" (Elsbach & Sutton, 1992) or "greenwash" (Ioannou et al., 2022) about their environmental, social, and governance (ESG) activities is a major concern for stakeholders (Baker et al., 2022; Bolino et al., 2008; Frankental, 2001; Hooghiemstra, 2000; Hopwood, 2009; Milne & Grubnic, 2011; Reitmaier et al., 2024). ESG reports are fertile ground for impression management because (1) related disclosures are difficult to substantiate and standardize, and (2) intended audiences extend beyond investors to stakeholders such as consumers, employees, and activists, who may be less financially and quantitatively oriented (Christensen et al., 2021). Mandatory ESG reporting frameworks (e.g., European Commission (2021)) seldom regulate the use of aesthetic images, which remain open to preparers' discretion and judgement.

In this paper, we analyze images using deep learning and artificial intelligence (AI) methods to examine the implications of image usage for disclosure quality in ESG reports. We (1) document the prevalence of image usage in ESG reports on a wide scale, (2) test the extent to which image usage is linked to firms' motivations for impression management, such as social disapproval and poor ESG performance, (3) contrast image usage that is likelier to be related to impression management against image usage that is likelier to be innocuous or even informative, and (4) explore whether stakeholder perceptions of firms are related to image usage.

Extensive research examines corporate impression management in numerical form (as reviewed in Dechow et al. (2010)) and narrative form (as reviewed in Loughran and McDonald (2016)). However, psychological experiments often find that pictures leave longer-lasting and

☆ The authors would like to thank workshop participants at Rutgers Business School, Southwestern University of Finance and Economics, the 2021 Hawai'i Accounting Research Conference, the 11th International Conference of The Japanese Accounting Review, the 2021 Baruch-Rutgers-Fordham Accounting Conference, and the 2021 Mid-Atlantic American Accounting Association Region Meeting. In particular, we thank discussants Zhongwen Fan and Danielle Green. We also appreciate valuable comments from Yu-an Chen, Ryan McDonough, Dan Palmon, Stephen Ryan, Bharat Sarath, and Han Yan. Keyi Zhao would like to thank her advisor Divya Anantharaman and dissertation committee members Feng Gao, Hussein Issa, and Hariom Manchiraju for their guidance. Divya Anantharaman and Disen Huang Albrecht would like to thank Rutgers University for funding. Keyi Zhao would like to thank Rutgers University as well as the Southwestern University of Finance and Economics for funding.

* Corresponding author.
** Co-corresponding author.
E-mail addresses: disen.huang@rutgers.edu (D.H. Albrecht), divyaa@rutgers.edu (D. Anantharaman), zhaoky@swufe.edu.cn (K. Zhao).

https://doi.org/10.1016/j.aos.2025.101616
Received 4 June 2022; Received in revised form 23 October 2025; Accepted 24 October 2025
Available online 10 November 2025

stronger impressions than text (Knobloch et al., 2003; Paivio, 1969, 1991; Zillmann et al., 1999), leading communications researchers to examine "visual rhetoric" (Courtis, 2004; Graves et al., 1996). Prior literature observes two main types of visual rhetoric: 1) data visualizations with numbers and text, such as graphs, charts, and, more recently, infographics; and 2) pictures without numbers or text, such as photographs and illustrations, which we term "aesthetic images" for differentiation and "images" for short. The literature on financial reporting visual impression management has focused on the first type (data visualizations). It establishes that in annual reports, companies strategically present graphs and charts to enhance perceptions of products, operations, credibility, and financial performance.[1] In contrast, the second type (aesthetic images) is relatively overlooked in empirical research. A handful of studies describe aesthetic images as conveying corporate missions and values, which often reflect societal trends (Benschop & Meihuizen, 2002; Graves et al., 1996; Preston et al., 1996).

The proliferation of ESG reports presents an exceptional opportunity to investigate image usage (see Appendix I for examples) in corporate communications. Cho et al. (2009) theorize that images enhance perceptions of ESG performance and commitment through "media richness" and "multiple cues." However, the difficulty of manual data collection has limited archival studies to case studies, single-industry studies, and small-sample studies (Cho et al., 2012; García-Sánchez and Araujo-Bernardo, 2019; Jones, 2011; McWilliams et al., 2006; Merkl-Davies & Brennan, 2007). Without widening samples and quantifying variations, our knowledge of how companies use and how stakeholders perceive aesthetic images is incomplete (Elsbach & Sutton, 1992). We leverage recent advances in AI—specifically, object detection with deep learning—to examine the large-scale determinants and consequences of aesthetic images in ESG reports, for 526 unique U.S. public companies' stand-alone ESG reports over the 14-year period between 2005 and 2018.

We formulate two measures of report-level image usage: the number of images scaled by the number of pages (IMAGES/PAGE, read as "images per page") and the number of images scaled by units of 1000 words (IMAGES/TEXT, read as "images to text"). Longitudinally, a growing pattern exists for reports' image usage for a given level of words, as shown in Fig. 1. Average IMAGES/TEXT exhibits a sharp upwards trend, strengthening the relevance and timeliness of rigorously examining image usage in ESG reports.

First, we examine whether image usage is associated with motivations for impression management identified in prior literature. We observe higher ESG report image usage in socially problematic industries that supply products or services that reinforce undesirable, unhealthy habits or that generate negative externalities (e.g., alcohol, tobacco, weapons). We also observe higher image usage in firms that do not commit to the Global Reporting Initiative (GRI), a set of voluntary ESG reporting standards representing "global best practice" (Global Reporting Initiative, 2020). Importantly, we find that firms with less extensive textual content and poor ESG performance exhibit higher image usage. We further find that heavier image usage holds predictive power for incidents of future ESG misconduct (e.g., environmental fines and discrimination lawsuits), incremental to that of ESG ratings. Collectively, these results suggest that image usage reflects firms' attempts at visual impression management.

However, some image usage may still serve innocuous or even informative functions. To identify the degree of strategic usage, we use a combination of AI and manual methods. We classify images into two categories: 1) firm-specific images that add substance or nuance (e.g., employees in company T-shirts picking up trash in their local community, energy-efficient fleets of electric trucks, resource-conserving infrastructure such as green buildings and recycling facilities, and earth-friendly products such as those with reduced or biodegradable packaging), and 2) generic or stock images that evoke ESG topics (e.g., a photo of a skyline with trees, oceans, animals, and modeled corporate meetings). We refer to the former as "specific images" and the latter as "generic images." As we hypothesize, generic image usage is more strongly associated with proxies for firms' motivations for impression management than is specific image usage.

We complete the analysis on image usage by assessing whether it influences stakeholders' perceptions of value. Whether visual impression management affects decision-making outside laboratory settings is a fundamental yet largely unanswered question (Beattie & Jones, 2008; Diouf & Boiral, 2017).[2] Experiments find that sophisticated audiences—those with financial and quantitative skills, training, or experience—discount embellished information (Elliott, 2006; Elliott et al., 2017; Frederickson & Miller, 2004). Surveys find that socially responsible investing (SRI) professionals such as fund managers, analysts, and consultants consider multiple information sources beyond ESG reports (Diouf & Boiral, 2017). We complement these studies by finding visual impression management to be associated with a returns-based proxy of perceived value among less sophisticated audiences like retail traders and individual shareholder activists, with this association being stronger for generic images than for specific images. Visual impression management is *not* associated with enhanced perception in institutional shareholders. We also find that ESG award committees reward specific image usage and penalize generic image usage. We conclude that experimental insights on the misleading effect of visual impression management translate into real-world impact, with the effects mitigated by stakeholders' sophistication and expertise.

We contribute to the literature on corporate disclosure, whose focus has expanded from numerical information to textual characteristics. Established topics include the linguistic complexity of filings (Li, 2008), the "walking down" of management guidance (Bartov et al., 2002), the tone of conference calls (Huang et al., 2014), and the salience of non-GAAP earnings (Bowen et al., 2005). We examine another prominent element of disclosures, images, which have been "neither widely nor systematically analyzed in social science research despite the richness of the information they contain" (Adukia et al., 2023). In mandatory disclosures, images reinforce textual or numerical information (Ben-Raphael et al., 2023; Christensen et al., 2024; Deng et al., 2023). In less regulated disclosures, images grab the scarce attention of social media users (e.g., earnings announcement Tweets in Nekrasov et al. (2021)), communicate forward-looking operational information (e.g., executive presentations in Cao et al. (2023)), and capture aggregate news sentiment (e.g., news photos in Obaid and Pukthuanthong (2022)). We document that images in ESG reports both communicate information and manage impressions.

In using AI techniques for image analysis, our paper relates to a burgeoning literature in economics and social science. Using similar techniques, recent studies examine minority and female representation in children's books (Adukia et al., 2023), style shifts in yearbook photographs (Voth & Yanagizawa-Drott, 2024), how profile pictures affect microlending decisions (Athey et al., 2023), and how mug shots affect sentencing decisions (Ludwig & Mullainathan, 2024).

[1] See, for example, Taylor & Anderson, 1986; Tufte, 1983; Steinbart, 1989; Beattie & Jones, 1992; Graves et al., 1996; Beattie & Jones, 2000a; Beattie & Jones, 2000b; Hancock, 2003; Godfrey et al., 2003; Mather et al., 2005; Penrose, 2008; Muino & Trombetta, 2009.

[2] Beattie and Jones (2008) characterize the literature on the consequences of corporate disclosure graph usage as "in its infancy," having only been "addressed using an experimental … approach" (P.75). The literature on the consequences of aesthetic image usage is yet less mature.

Diouf and Boiral (2017) fill a gap in the survey literature by interviewing practitioners who regularly use ESG information rather than business leaders who may be removed from the implementation of ESG assessment. Our empirical evidence corroborates, on a large scale, such survey evidence that sophisticated stakeholders value and use information beyond ESG reports.

Further, in examining aesthetic images, we fill a gap in the fast-grown ESG disclosure literature by examining a presentation choice (Andrew & Baker, 2020). The voluntary disclosure literature faces a challenge in disentangling choices in reporting from choices in activities (Christensen et al., 2021). By distinguishing generic images (which reflect choices in presentation and reporting) from specific images (which reflect underlying ESG activities), we shed light on the decoupling of reporting and activities. Further, we contribute archival evidence to the growing debate on accountability in ESG reporting.[3] Our findings carry meaningful implications for practitioners and regulators as well; to the extent that stakeholder perception affects decisions, visual impression management affects the assessment of corporate externalities and allocation of resources (Chatterji et al., 2009).

Finally, by examining the underexplored impact of organizational impression management on markets (Beattie & Jones, 2008; Hopwood, 2009), we contribute to the impression management literature at the intersection of psychology, organizational behavior, strategy, management, and behavioral economics. Our results confirm some of the observations from experimental and small-scale studies, but caution against over-generalization to wider market participants. Existing studies and review papers focus on the possibility that visual impression management has positive impact or no impact. We document the first evidence of which we are aware that visual impression management elicits negative reactions from expert audiences such as award committees. We invite more researchers to shed more light on the mechanisms and conditions of this phenomenon using archival, field, and experimental methods.

The rest of this paper unfolds as follows. Section 2 summarizes background literature, develops the hypotheses, and describes our research design. Section 3 describes our image recognition and validation method and our sample. Sections 4 and 5 present empirical results on determinants and consequences of image usage, respectively. Section 6 concludes and discusses opportunities for future research.

2. Hypothesis development and research design

2.1. Legitimacy theory and impression management

In legitimacy theory, organizations ensure survival by showing that the values expressed through their activities align with the values of society (Dowling & Pfeffer, 1975). Jensen and Meckling (1976) assign shareholders primacy in granting corporate legitimacy, while stakeholder theory also considers customers, suppliers, employees, surrounding communities, governments, and others impacted by corporate activities (Donaldson & Preston, 1995; Freeman, 1984). Firms can increase legitimacy in stakeholders' eyes through corporate ESG, which consists of "actions that *appear* to further some social good, beyond the interests of the firm and that which is required by law" (McWilliams & Siegel, 2001, p. 117; emphasis added). In addition, to build and defend legitimacy, firms often communicate their actions to stakeholders (Dowling & Pfeffer, 1975). While some forms of impression management are likely ethical, firm communications that mislead stakeholders potentially are unethical and/or illegal forms of impression management.

Fig. 2 diagrams the conceptual role of impression management in unduly influencing stakeholder perception. If stakeholders perceive ESG performance as high, firms enjoy numerous potential benefits: enhanced financial performance and firm value (Al-Tuwaijri et al., 2004; Elliott et al., 2014; Flammer, 2015a; Lev et al., 2010), lower idiosyncratic risk (Bansal & Clelland, 2004; Neu et al., 1998), lower litigation risk (Koh et al., 2014), lower cost of equity (Dhaliwal et al., 2011), lower cost of debt (Cooper & Uzun, 2015), and protection against value declines during unfavorable events (Godfrey et al., 2009; Lins et al., 2017; Minor & Morgan, 2011).

Perceptions of ESG are susceptible to impression management for several reasons. First, ESG disclosures deal with long-term prospects and intangible constructs (e.g., customer goodwill and employee loyalty), which are difficult to quantify (Christensen et al., 2021). Second, ESG outputs may not be fully expressible in financial terms, leading observers to disagree on their measurement (Kitzmueller & Shimshack, 2012). Third, expressions of ESG commitment can increase organizational legitimacy, even in the absence of substantiating activities (Bansal & Clelland, 2004; Dowling & Pfeffer, 1975; Hopwood, 2009).

2.2. Visual impression management

As we summarized above, firms may engage in various forms of impression management to enjoy the benefits of high ESG perception without incurring the costs of impactful ESG activities. The direct costs of executing visual impression management are low. A professionally produced ESG report package costs thousands to tens of thousands of dollars (Graphic Artists Guild, 2021). After firms have contracted with a graphic designer (a common practice), the incremental cost of generating visual appeal is also low. However, users process visual information more subconsciously than textual or numerical information (Paivio, 1991). Visual impression management is thus a double-edged sword for companies, in that its costs and risks are low, but its benefits are unpredictable.

Firms plausibly expect that visual impression management carry lower indirect costs as well. These costs stem from the risk of detection, including reputational losses, regulatory enforcement, and other negative consequences. Although stakeholders react negatively to deception in financial and nonfinancial settings (Barton & Mercer, 2005; Ioannou et al., 2022; Pacheco-Ortiz et al., 2024), narrative and numerical inconsistencies are more readily verifiable than aesthetic judgments, which are inherently more subjective. Even if stakeholders could pinpoint how visuals enhance their perception, they could hardly cite aesthetic images as evidence in disputes. In contrast, narrative and numerical information supply evidence when a firm's actions or results do not match its promises or explanations. For example, SEC Climate and ESG Task Force held BNY Mellon accountable for falsely claiming its funds had undergone ESG evaluations (SEC, 2022), illustrating an enforcement cost for narrative impression management. Overall, we expect visual impression management to be one of the least costly forms of impression management, both directly and indirectly. Such low costs make it more likely that net benefits are positive.

If the cost of visual impression management is so low, why do firms not use even more images than we observe? One reason is that the nature and purpose of aesthetic images impose a natural constraint. Aesthetics are meant to accompany core content (i.e., text and numbers) and cannot stand alone. Another possibility is that image usage is subject to diminishing returns. Other factors, like industry norms, design trends, and firm preferences, also contribute to variation in image usage. Some baseline image usage may be expected. Our goal is to separate expected from unexpected image usage, and then test for strategic motivations behind the latter. We achieve this by 1) scaling image measures by report length and words, as outlined in Section 1; and 2) controlling for firm characteristics and/or industry characteristics as appropriate in multivariate regressions (see details in Sections 2 and 4). We hypothesize in the next section that holding these other factors constant, ESG report image usage is affected by the expected benefits of and constraints

[3] ESG information potentially leads to financial effects (Elliott et al., 2014; Ioannou & Serafeim, 2015; Lys et al., 2015; Martin & Moser, 2016; Manchiraju & Rajgopal, 2017; Grewal et al., 2019; Bartov et al., 2002; Desai et al., 2023; see also Fig. 2 in Section 2), real effects (Christensen et al., 2017; Darendeli et al., 2022), or no effects (Aswani & Rajgopal, 2023; Larcker & Watts, 2020; Raghunandan & Rajgopal, 2023). ESG information also serves as public relations and advertising (as reviewed in Shabana and Ravlin (2016)). A contemporaneous working paper, Azevedo-Rezende et al. (2021), studies green pixels in ESG reports as a tool for greenwashing.

on visual impression management. For clarity and conciseness, we present the constructs in each hypothesis together with the proxies we use to test it.

2.3. Motivations and opportunities for impression management

We begin by examining firm characteristics which serve as motivations for impression management by increasing the expected benefits. Prior literature suggests that the need for positive ESG perception varies by industry (Hooghiemstra, 2000). In this hypothesis, we examine industry membership as an explanatory variable.

First, we identify industries with social stigma. We follow (1) Spiller (2000), Fabozzi et al. (2008), and Hong and Kacperczyk (2009) to classify alcohol, tobacco, and gaming as sin industries, (2) Dupire and M'Zali (2018) to define manufacturing, mining, and chemicals as dirty industries, and (3) Cai et al. (2012) to define weapons, oil, cement, and biotech as controversial industries. We refer to sin, dirty, and controversial industries collectively as problematic industries. Second, we consider industries where firms strategically use ESG to achieve product differentiation and higher revenues (Flammer, 2015b; Dupire & M'Zali, 2018). We use the Hoberg and Phillips (2016) textual analysis measure to identify highly differentiated industries. Third, we consider industries with high sensitivity to consumer perception, such as business-to-consumer (B2C) industries (Bagnoli & Watts, 2003; Flammer, 2015b; Lev et al., 2010). We define B2C industries as consumer goods and finance, following Lev et al. (2010). Our hypothesis follows:

Hypothesis 1. All else equal, industries expecting higher benefits from positive ESG perception (i.e., sin, dirty, controversial, highly differentiated, or B2C industries) exhibit higher image usage.

However, some firms may be unable to engage in visual impression management to the extent predicted by their motivations. We observe that GRI reports appear text-heavy, perhaps to meet content requirements while keeping page numbers reasonable. While Moneva et al. (2006) and Boiral (2013) caution that GRI adoption does not guarantee compliance, Schadewitz and Niskala (2010) and Reverte (2012) link GRI adoption to higher firm value and lower cost of capital. We posit that GRI standards constrain firms in report content and presentation, increasing spatial costs and reducing opportunities for impression management. From this hypothesis onwards, we use industry membership as a control variable to account for expected image usage driven by industry-specific factors.

Hypothesis 2. All else equal, firms with more constraints (i.e., GRI Standards) exhibit lower image usage in ESG reports.

2.4. Poor ESG performance as motivation for impression management

We hypothesize that poor ESG performance is a key motivation for visual impression management, as embellishing ESG disclosure is less costly than improving ESG performance. However, negative ESG performance cannot be hidden indefinitely. When facts eventually come to light, firms attempting to obscure poor ESG performance with visual impression management should exhibit negative consequences. We formulate this hypothesis in two related parts:

Hypothesis 3(a). All else equal, companies with poor ESG performance so far exhibit higher image usage in ESG reports.

Hypothesis 3(b). All else equal (including ESG performance), higher image usage in ESG reports is associated with more incidents of ESG misconduct in the future.

To test Hypothesis 3(a), we use ESG ratings to proxy for current ESG performance. Namely, we use the two component scores underlying Refinitiv's overall ESG performance assessment—the ESG Score, which is based on corporate reports, and the ESG Controversies Score, which is based on external, media-published sources (Refinitiv, 2019).[4] The ESG Score summarizes Refinitiv's assessment of the firm's professed ESG policies, initiatives, and activities. The ESG Controversies Score functions as a rather extreme measure of revealed negative ESG performance (Aouadi & Marsat, 2018; Dorfleitner et al., 2020). Both scores are indispensable, as firm-produced reports often do not include negative ESG news (Diouf & Boiral, 2017). We define SCORE_ESG as equal to the Refinitiv ESG Score and SCORE_CONTROVERSIES as 100 – Refinitiv ESG Controversy Score (so that a higher SCORE_CONTROVERSIES intuitively reflects worse performance).

To test Hypothesis 3(b), we turn to fundamental indications of poor ESG performance as an outcome variable, in the spirit of Chatterji et al. (2009). We do not use future ESG ratings, which are autocorrelated with current ESG ratings. We gather records of corporate misconduct from Violation Tracker, a nonprofit database of corporate harm represented by regulatory penalties and class action lawsuit settlements, with an emphasis on environmental, health, and safety issues (Good Jobs First, 2015). It is used in increasing numbers of ESG studies to track corporate misconduct, e.g., Soltes (2019), Heese et al. (2022), and Raghunandan and Rajgopal (2023).

In an archival setting, identifying impression management directly is challenging because (1) we can only make inferences from observed image usage, and (2) what firms intend to achieve with images and what stakeholder perception would have been without those images are both unobservable. Thus, in Hypotheses 1–3, we compile an array of predictions that higher image usage aligns with firms' motivations and opportunities for impression management.

2.5. Impact of ESG report image usage

We investigate the impact of image usage on stakeholder groups who are likely to consider ESG information in decision-making, consume ESG reports while gathering ESG information, and exhibit large-scale, measurable responses. We arrive at two major stakeholder groups: 1) equity investors, who remain major stakeholders (Jensen, 2002), especially as interest in SRI has grown, and 2) ESG award committees. The ESG award setting also presents a rare opportunity to widen our stakeholder tests to "third sector" organizations (Cerrato & Ferrando, 2020), including nonprofits, non-governmental organizations, consumer associations, and community groups.

We organize perception tests roughly by stakeholder sophistication, as experimental research finds that more sophisticated audiences—those with greater training, knowledge, or experience in business matters—are more likely to "see through" impression management than unsophisticated audiences (Elliott, 2006; Elliott et al., 2017; Frederickson & Miller, 2004). We posit that unsophisticated users (e.g., retail traders) are more susceptible to visual impression management than sophisticated users (e.g., institutional shareholders and ESG award committees). First, we examine the reactions of general shareholders, as ESG considerations affect the demand for and price of shares (Chatterji et al., 2009; Li et al., 2023). We expect to observe negative future cumulative abnormal returns (CAR) in association with visual impression management if 1) ESG performance is value-relevant, 2) investors are susceptible and overvalue firms with strategic image usage at ESG report publication, and 3) return reversals occur as communications become

[4] Refinitiv (formerly Thomson Reuters Financial & Risk and now owned by the London Stock Exchange Group (LSEG)) scores are the same as the Thomson Reuters ESG scores in earlier literature, which Thomson Reuters acquired from Asset4 (Thomson Reuters, 2018). Under these different names, the scores have been used in 285 empirical ESG studies published in business journals between 2011 and 2021 (De Villiers et al., 2022). We find qualitatively similar relationships between ESG performance and ESG report image usage when using another popular proxy for ESG performance, MSCI KLD ratings (see Online Appendix, Table OA4).

alternately verified or invalidated over time. Second, we investigate the reaction of activist shareholders (i.e., whether their satisfaction increases with visual impression management). Third, we investigate the reaction of ESG award committees (i.e., whether their propensity to give awards increases with visual impression management).

Hypothesis 4. All else equal, higher image usage in ESG reports is associated with positive stakeholder perception (i.e., lower CAR due to later return reversals, fewer shareholder proposals, and a higher likelihood of winning ESG awards). We expect this positive association to be more pronounced in less sophisticated stakeholders.

2.6. Information content of images

Interwoven with the previous hypotheses, we consider the possibility that some aesthetic images are more informative than others. This approach coincides with that of three contemporaneous and independently developed papers in the annual report (Ben-Raphael et al., 2023; Deng et al., 2023) and crowdsourcing settings (Ronen et al., 2023). While it is common for corporate reports to use generic images (e.g., Exhibit 4 in Appendix I), many graphic design packages involve a dedicated photographer who takes firm-specific photos (Herring, 1990). The recent proliferation of phone cameras and social media has increased the supply of firm-specific photos. Unlike generic images, specific images contain substantiating information (e.g., photos of energy-efficient products).

Since specific images require greater customization, they are generally more costly for ESG impression management purposes. Also, specific images may incur other costs, such as the proprietary cost of revealing innovation to competitors (Verrecchia, 1983). Thus, impression management concerns are more severe for generic images than for specific images. We contrast specific images against generic images throughout the posited relationships in Hypotheses 1–4:

Hypothesis 5. Motivations and effects of impression management manifest less in specific images than in generic images.

We classify images as specific or generic using a combination of automated and manual methods, detailed in Section 3 below. Note that specific images can appropriately enhance perception by substantiating information or inappropriately enhance perception by generating positive feelings with limited basis in a firm's fundamentals; we cannot disentangle these effects. However, we still expect generic images to contain more visual impression management elements than specific images. Our empirical strategy is not to estimate the absolute relationship of specific or generic images to impression management, but to draw inferences by contrasting these two relationships.

In the next section, we turn to the measurement of images.

3. Image measurement and data

We manually download U.S. public companies' ESG reports from 2005 to 2018 in the Sustainability Disclosure Database, recording the fiscal year covered by each report from its title page. To fill gaps in the database, we conduct additional searches through Google and Corporate Register, yielding a total of 3423 ESG reports in Adobe PDF format. We also manually collect the publication month and year of each report from Corporate Register. We exclude observations missing control variables for financials or textual disclosure quality. We merge these ESG reports by firm-year with Compustat for firm financial characteristics and with the Bloomberg Disclosure Score database for report content, resulting in 2602 reports for further analysis.

We convert the PDF reports into Microsoft Word documents using two types of conversion software. In rare cases where conversion fails, we eliminate the reports from the sample. Next, using Python, we convert the Word files into ZIP folders and delete, from the folders, any image below 20 KB in size. These very small images are mostly design elements such as lines, logos, and banners. We process the remaining 133,431 images through Faster R–CNN, a state-of-the-art deep learning network for image object detection that has been extensively validated in computer science research. Object detection reduces Type II errors in identifying aesthetic images because images without detectable objects are likely to be data visualizations. Faster R–CNN labels each detected object with its detected category (e.g., human, plant, building) and a confidence score between 0 and 1. We keep 81,982 images with objects detected at higher than 0.5 confidence, the more-likely-than-not threshold.

Each of the three co-authors then independently manually samples 1 % of each category or 20 images, whichever is greater, for a total of 1610 images across 62 categories. Please see details of the manual sampling schema in Appendix II. We classify images as specific if the detected objects display brands or names or otherwise appear personalized, according to our best judgment. We delete eight high-inaccuracy categories (e.g., those labeled as "candles" or "flags," which we found to be mostly bars and backdrops used in charts). We reconcile independent sampling by discussion to reach consensus, resulting in a sample of 69,839 object-detected images and 66,308 images with specificity/generality weights. Overall, we find that larger images tend to be less specific (see Appendix II Section 2.2.2). Table 1 Panel A summarizes the sample construction.

Table 1 Panel B presents average image usage by industry, following the Fama and French (1997) 12 industry classification. In Columns 1&2, Consumer Non-Durables exhibits the highest average image usage, while Business Equipment exhibits the lowest. This is preliminarily consistent with B2C industries being more sensitive to consumer perception than B2B industries. Comparing Columns 3–4 against Columns 5–6, we observe that the average usage of specific images exceeds that of generic images in every industry. In fact, specific image usage is nearly twice as high as generic image usage. This finding is surprising, given the sizable stream of prior literature that perceives ESG reports as performative (Shabana & Ravlin, 2016).

We summarize industry insights from another perspective in Appendix II Section 2.2.3. With two exceptions, heavy manufacturing industries tend to lag in ESG report image specificity. The exceptions, surprisingly, are Consumer Durables and Oil, Gas, and Coal, which lead in image specificity by showcasing tangible assets. Service industries like Finance and Healthcare also lead in image specificity, but through showcasing people. However, if we hold firm characteristics constant, the Finance sector uses fewer images than other industries (untabulated). These descriptive results highlight the complexities of ESG report image choices and the limitations of univariate explanations.

Table 1 Panel C presents summary statistics for all variables. We obtain stock prices from CRSP, analyst following from I/B/E/S, institutional ownership from Thomson Reuters 13-F, press releases from Capital IQ, shareholder proposals and mutual fund voting from ISS Voting Analytics, and ESG awards from Refinitiv. Sixty-six percent of sample reports declare that they follow GRI Standards.

We note that images appear in almost every sample report. The median number of IMAGES/PAGE is 0.53, for a median report length of 45 pages. To control for baseline graphic design preferences, we collect annual report image usage in the same firm-year. ESG report image usage has grown during the sample period, while annual report image usage has fallen from its peak in the mid-2000s. Average image usage is now higher in ESG reports than in annual reports.

Table 1 Panel D presents abbreviated pairwise correlations between aesthetic images, firm characteristics, and outcome variables. The correlation between IMAGES/PAGE and IMAGES/TEXT is high at 0.81 and significant at the 0.1 level. Although we expect image usage to show persistence, the correlation between current and lagged IMAGES/PAGE (IMAGES/TEXT) is only 0.51 (0.56). In untabulated analyses, after controlling for textual content quality, firm characteristics, and industry and year fixed effects, lagged image usage is still significantly positively associated with current image usage at the 1 % level, with a coefficient of 0.4453 (0.4777) for IMAGES/PAGE (IMAGES/TEXT). The correlation between ESG report and annual report image usage is close to 0 and statistically insignificant. ESG report image usage correlates negatively with TEXT_CONTENT, indicating that image-heavy reports disclose fewer topics.

Correlations between the three types of image usage (overall, specific, and generic) are reported in Appendix II. The correlations for IMAGES/PAGE (IMAGES/TEXT) range from 0.794 to 0.916 (0.837–0.926). Thus, it is critical to control for generic image usage when analyzing specific image usage and vice versa. Note that specific image usage and generic image usage are related by construction but not multicollinear: theoretically, if we hold total image usage constant, specific image usage and generic image usage should be negatively correlated; empirically, however, reports that use more specific images also use more generic images, resulting in higher image usage overall.

We report the results of hypothesis testing in Sections IV and V.

4. Results on determinants of ESG report image usage

4.1. Report and firm characteristics as determinants of image usage

We describe how ESG report image usage varies with disclosure quality and firm characteristics using the following model:

$$\text{Image Usage}_{i,t} = \alpha + \beta_1\text{TEXT\_CONTENT}_{i,t} + \beta_2\text{TEXT\_SPECIFICITY}_{i,t} + \beta_3\text{AR\_Image Usage}_{i,t} + \beta_4\text{MKTCAP}_{i,t} + \beta_5\text{AGE}_{i,t} + \beta_6\text{BTM}_{i,t} + \beta_7\text{LEV}_{i,t} + \beta_8\text{ADVT} + \beta_9\text{RD}_{i,t} + \beta_{10}\text{CAPX}_{i,t} + \beta_{11}\text{ROA}_{i,t} + \beta_{12}\text{SALES}_{i,t} + \beta_{13}\text{SALESGROWTH}_{i,t} + \beta_{14}\text{TURNOVER}_{i,t} + \beta_{15}\text{INSTHOLDINGS}_{i,t} + \beta_{16}\text{ANALYSTS}_{i,t} + \beta_{17}\text{NPR}_{i,t} + \beta\text{Industry Fixed Effects}_i + \beta\text{Year Fixed Effects}_t + \varepsilon$$

$$(\text{Where Image Usage}_{i,t} = \text{IMAGES/PAGE}_{i,t} \text{ or IMAGES/TEXT}_{i,t}) \quad (1)$$

We focus on two aspects of ESG report disclosure quality: extensiveness and specificity. Companies can use as many or as few images as they like while issuing high-quality or low-quality reports. However, if companies attempt visual impression management as a less costly alternative to improving disclosure content, then image usage may be negatively associated with textual content quality. To proxy for disclosure extensiveness (TEXT_CONTENT), we use the Bloomberg ESG Disclosure Score, which captures the substance of disclosure content without regard to presentation choices. To produce this score, Bloomberg maintains a list of important ESG topics and tracks the number of such topics in each ESG report. Many empirical studies use this score to measure the extensiveness or comprehensiveness of ESG disclosures, e.g., Li et al. (2018) and Christensen et al. (2022). We also create a composite measure of linguistic specificity (TEXT_SPECIFICITY) that increases with specific language (Finkel et al., 2005; Hope et al., 2016) and decreases with boilerplate language (Lang & Stice-Lawrence, 2015). In these previous studies, higher specificity is associated with lower narrative impression management.

To examine firm determinants of ESG report image usage, we implement McWilliams and Siegel (2000)'s suggestion to highlight two ESG-relevant financial characteristics: 1) advertising expense (ADVT), related to firms' effort to bring ESG initiatives to the attention of consumers, and 2) research and development expense (RD), related to firms' use of resources in ESG innovation. We further control for the following financial characteristics: log transformed market capitalization (MKTCAP), log transformed firm age (AGE), book-to-market (BTM), leverage (LEV), capital expenditures (CAPX), return on assets (ROA), sales revenues (SALES), and year-on-year growth in revenues (SALESGROWTH). Larger firms are more visible (Dowling & Pfeffer, 1975), and higher capital expenditures may involve higher emissions or pollution, increasing motivations for impression management. To proxy for a firm's liquidity and information environment, we use share turnover (TURNOVER), institutional ownership (INSTHOLDINGS), and analyst following (ANALYSTS). To proxy for the intensity of the firm's voluntary disclosure, we follow Kirk and Vincent (2014) and use the number of press releases published that year (NPR). Due to lack of theoretical guidance, we refrain from making directional predictions on the relationships between financial characteristics and image usage.

Table 2 presents the ordinary least squares (OLS) estimation results for Model 1. The coefficient on TEXT_CONTENT is negative and significant at the 1 % level in all specifications, indicating that ESG report image usage is higher when textual content quality is lower. A one-standard-deviation decrease in TEXT_CONTENT is associated with 29.6 % of a one standard deviation increase in IMAGES/PAGE or 35.4 % of a one standard deviation increase in IMAGES/TEXT. Interestingly, this negative relationship is found both in specific images (Columns 3–4) and generic images (Columns 5–6). It is *incrementally* significant at the 1 % and 5 % levels, respectively, even when we control for the other type of image usage. Thus, images could be used as substitutes for textual disclosure. However, image usage is not antithetical to *all* aspects of disclosure quality: the coefficient on TEXT_SPECIFICITY is positive and significant, indicating that ESG reports using more images exhibit higher linguistic specificity. One possible explanation is that firms that expend more effort on ESG reports use more images *and* disclose more specific information; another is that firms treat narrative and visual impression management as substitutes. Future research is needed to untangle these explanations.

Annual report image usage holds no explanatory power for ESG report image usage, supporting the view ESG reporting and financial reporting are driven by different considerations. With a few exceptions (e.g., the positive association between IMAGES/TEXT and institutional ownership), financial characteristics also hold low explanatory power.

Interestingly, financial characteristics exhibit different relationships with specific and generic image usage. For example, larger firms exhibit heavier use of specific images and lower use of generic images, as evidenced by the positive (negative) coefficients on MKTCAP in Columns 3–4 (5–6). This difference is consistent with larger firms having more resources to order customized photos rather than license generic photos. In contrast, we find that R&D-intensive firms use more generic images and fewer specific images. We conjecture that firms may not wish to show asset innovations in specific images due to concerns about proprietary costs (Verrecchia, 1983). Finally, institutional ownership shows stronger associations with generic images than specific images, although both associations are positive.

Overall, we find nuanced relationships between aesthetic images and textual characteristics in ESG reports: images are negatively related to content extensiveness but positively related to textual specificity. These nuanced relationships suggest that innocuous or informative image usage may be present alongside strategic image usage. We continue to unpack strategic image usage in the next section.

4.2. Industry characteristics and GRI adoption

To test Hypothesis 2, we add Industry $\text{Group}_{i,t}$ (alternately set to $\text{SIN\_IND}_{i,t}$, $\text{DIRTY\_IND}_{i,t}$, $\text{CONTROVERSIAL\_IND}_{i,t}$, $\text{B2C\_IND}_{i,t}$, or $\text{HIGHDIFF\_IND}_{i,t}$) to the independent variables in Model 1 and exclude industry fixed effects to avoid subsuming industry characteristics.[5] In Table 3 Columns 1–10, eight out of the ten coefficients on industry groups predicted to have high motivations for impression management are positive and significant. That is, firms operating in sin, dirty, controversial, highly differentiated, or B2C industries on average use more images, in line with Hypothesis 2. For a report of average length (55.38 pages or 18.70 thousand words), impression management motivated industries would use 2–6 more images.

Next, we add $\text{GRI}_{i,t}$ to Model 1 and report the estimation results in Table 3 Columns 11–12. The coefficients on GRI are significantly

[5] We conduct related univariate testing in Table OA1 Panels A-E of the Online Appendix, where results are generally consistent with multivariate results, with the exception of controversial industries.

negative, indicating that on average, GRI reports exhibit lower image usage than non-GRI reports. For a report of average length, GRI adoption is linked to 5–8 fewer images. Further, the lower usage is more evident in generic images than specific images (Table OA2 Panels A vs. Panel B in the Online Appendix). These results support Hypothesis 2 that GRI adoption constrains visual impression management.

4.3. ESG performance

In this subsection, we test whether poor ESG performance motivates firms to attempt visual impression management.

4.3.1. Existing ESG performance

Recall that SCORE_ESG proxies for comprehensive aspects of ESG performance from firm-reported sources and SCORE_CONTROVERSIES proxies for negative aspects of ESG performance from media-reported sources. In univariate comparisons, groups with lower SCORE_ESG tend to use significantly more images, but so do groups with lower SCORE_CONTROVERSIES. This univariate pattern is unsurprising given 1) the significantly positive correlation between SCORE_ESG and SCORE_CONTROVERSIES in Table 1 Panel D, and 2) findings in prior literature that firms rated high in ESG strengths tend to also rate high in ESG weaknesses (Chatterji et al., 2009). However, it seemingly presents a contradiction and highlights the importance of multivariate testing for Hypothesis 3(a). We use the following model:

$$\text{Image Usage}_{i,t} = \alpha + \beta_1 \text{ QUARTILE\_SCORE\_ESG}_{i,t} + \beta_2\text{SCORE\_CONTROVERSIES}_{i,t} + \beta_3\text{TEXT\_CONTENT}_{i,t} + \beta_4\text{TEXT\_SPECIFICITY}_{i,t} + \beta\text{Firm Controls}_{i,t} + \beta\text{Industry Fixed Effects}_i + \beta\text{Year Fixed Effects}_t + \varepsilon \quad (2)$$

In Table 4 Panel A, we divide SCORE_ESG into yearly quartiles and estimate Model 2. The bottom quartile of SCORE_ESG (i.e., firms with the poorest ESG performance) exhibits the heaviest image usage, the two middle quartiles a little less, and the top quartile the least. For a report of average length measured by pages (words), a firm in the bottom quartile of ESG performance uses 1.42 (3.48) more images than one in the third quartile, 1.41 (2.84) more images than one in the second quartile, and 4.88 (8.79) more images than one in the top quartile. The negative relationship between SCORE_ESG and image usage is more pronounced for generic images (Columns 5–6) than for specific images (Columns 3–4). The stronger association between generic images, which lack substance, and poor ESG performance suggests that poor-performing firms strategically use images to attempt visual impression management.

Note that though we control for ESG report textual characteristics, the coefficient on TEXT_CONTENT remains significantly negative. This result is consistent with past findings that firms with poor ESG performance tend to have less extensive, less quantifiable disclosures (Al-Tuwaijri et al., 2004). We show that the relationship between ESG performance and image usage is incremental to the relationship between ESG performance and textual disclosures.

Alternative specifications in Table OA3 of the Online Appendix reinforce the strong evidence that poor ESG performance (lower SCORE_ESG) is associated with greater image usage, and show modest evidence that media-reported controversies (SCORE_CONTROVERSIES) are also associated with greater image usage. We conclude that worse-performing firms use more images, especially generic images, which supports Hypothesis 3(a).

4.3.2. Future ESG performance

If firms use visual impression management to delay stakeholder perception of poor ESG performance, then excessive image usage should predict poor ESG performance in the future. To test this prediction, we use the natural log of the number of corporate misconduct incidents from Violation Tracker, MISCONDUCT, as the outcome variable. We transform SCORE_ESG into POOR_ESG = 100-SCORE_ESG so that it intuitively increases with poor performance. Our findings in Section 4.3.1 suggest that image usage is motivated by past ESG performance. Therefore, we interact image usage and past ESG performance in outcome tests:

$$\text{Outcome}_{i,t+1} = \alpha + \beta_1\text{Image Usage}_{i,t} + \beta_2\text{Image Usage}_{i,t} * \text{POOR\_ESG}_{i,t} + \beta_3\text{Image Usage}_{i,t} * \text{SCORE\_CONTROVERSIES}_{i,t} + \beta_4\text{POOR\_ESG}_{i,t} + \beta_5\text{SCORE\_CONTROVERSIES}_{i,t} + \beta_6\text{TEXT\_CONTENT}_{i,t} + \beta_7\text{TEXT\_SPECIFICITY}_{i,t} + \text{Firm Controls}_{i,t} + \text{Industry Fixed Effects}_i + \text{Year Fixed Effects}_t + \varepsilon \quad (3)$$

Table 4 Panel B Columns 1–2 report estimation results using all object-detected images. The coefficients on IMAGES/PAGE and IMAGES/TEXT are positive and significant, indicating that high image usage predicts future incidents of misconduct. Unexpectedly, the coefficient on the interaction between image usage and POOR_ESG is negative, indicating that POOR_ESG slightly weakens the power of image usage to predict misconduct. However, the coefficients on the interactions are not economically significant (at nearly one hundred times smaller than the coefficients on image usage). SCORE_CONTROVERSIES is a positive and significant predictor of misconduct, while POOR_ESG is not; this echoes Chatterji et al. (2009)'s findings on the predictive power of environmental weaknesses and strengths. A one-standard-deviation increase in IMAGES/PAGE (IMAGES/TEXT) is associated with 22.5 % (18.5 %) additional incidents of misconduct, or slightly less if the current ESG Score is low. Therefore, even after controlling for existing ESG performance, we consistently find that image usage predicts poor future ESG performance, supporting Hypothesis 3 (b).

Columns 3–4 (5–6) show the results for specific (generic) image usage. Specific image usage is significantly and incrementally associated with future MISCONDUCT, suggesting that we cannot rule out impression management motivations behind substantive images. Nevertheless, the coefficient on generic IMAGES/PAGE (Column 5) is twice the magnitude of the coefficient on specific IMAGES/PAGE (Column 3). Relatedly, although the coefficient on generic IMAGES/TEXT (Column 6) is not incrementally significant after we control for specific IMAGES/TEXT, it is still nearly double the magnitude of the coefficient on specific IMAGES/TEXT (Column 4). We cautiously conclude that generic image usage is associated with more incidents of future misconduct.

5. Results on stakeholder decisions in relation to ESG report image usage

Whether ESG report visual impression management affects stakeholder decisions outside of experimental settings is an open question for three reasons. First, real-world stakeholders may be more sophisticated and less susceptible than experiment participants, who are usually students. Second, in a world where news competes for limited attention (Christensen et al., 2017), stakeholders may not pay close enough attention to ESG reports to become susceptible. Third, real-world stakeholders, especially sophisticated institutions, routinely seek external sources of ESG information, such as media outlets and third-party ratings (Diouf & Boiral, 2017; SASB Symposium Roundtable, 2016). Any of these empirical considerations may prevent the replication of experimental results in archival settings.

Fig. 3 maps the stakeholder groups whose reactions we test to their empirical settings. We organize our settings to roughly increase in sophistication and decrease in susceptibility to visual impression management: individual shareholders, institutional shareholders, and ESG award committees.

5.1. Equity valuation

We calculate one-year-ahead CAR as the excess of realized returns over Fama-French 3-Factor returns for all December 31 year-end firms in the 12 months after ESG report publication. We run OLS regressions on

Model 3, with CAR as the outcome variable. As visual impression management may produce stronger impact in less sophisticated audiences, we expect stronger evidence of return reversals in firms with higher retail investor ownership (i.e., lower institutional ownership).[6]

We report estimation partitioned by institutional ownership in Table 5. (The average effect across all investors is insignificantly different from zero, which we do not tabulate for conciseness.) In Columns 7–12, the low institutional ownership subsample, the coefficient on IMAGE_USAGE*SCORE_CONTROVERSIES is significantly negative in all specifications, and the coefficient on IMAGE_USAGE*POOR_ESG is significantly negative in half the specifications. These coefficients indicate initial overvaluation and subsequent return reversals for poor-ESG firms with heavy image usage. However, the coefficients on image usage itself are significantly positive and larger—unconditional on ESG performance, unsophisticated investors in fact appear to be undervaluing the information content of images.

Multiple explanations are possible for the lack of statistically significant outcomes in the high institutional ownership subsample (Columns 1–6): 1) certain investors do not consider ESG performance in decision-making, 2) certain investors obtain ESG information elsewhere and do not rely on ESG reports, and 3) certain investors read ESG reports but are not affected by aesthetic images. We cannot disentangle these explanations using our current research design and look to future research to do so. Overall, we conclude that visual impression management is associated with overvaluation and return reversals only in retail investors.

5.2. Shareholder activism

In our second shareholder perception test, we investigate whether ESG report image usage is related to institutional and individual shareholder activism. Our sample institutional shareholder activists include for-profit funds and corporations but exclude labor unions and religious organizations, whose priorities are difficult to predict (Goodman, 2015; Yermack, 2010).[7] Individual activists include a few prominent professionals but many more unsophisticated "gadflies" (Gantchev & Gianetti, 2020). We therefore expect institutional shareholder activists to be less susceptible to visual impression management than individual shareholder activists, on average.

We count the number of shareholder proposals raised in year t+2 and construct its natural log as PROPOSALS. Since ESG reports pertaining to year t are published in year t+1, often after the proposal deadline (120 days before the annual meeting), they can only affect new proposals raised in year t+2. Around half of our sample receive at least one proposal, totaling 3740 proposals. Table 6 reports the estimation results on all object-detected images, partitioned by the sophistication of the proposal sponsor. In all specifications, PROPOSALS is significantly positively related to SCORE_CONTROVERSIES at the 1 % level, indicating that activist shareholders submit more proposals when companies experience more negative ESG news. The interaction term between image usage and SCORE_CONTROVERSIES is negative, suggesting that images soften activists' responses to negative ESG news. The softening effect on individual activists is driven by generic images (Panel C) rather than specific images (Panel B), suggesting that they are susceptible to impression management.

However, for sophisticated institutional activists, a different effect dominates the softening effect by orders of magnitude. The coefficient on image usage itself is significantly *positive* and 63–74 times as large as the coefficient on the interaction term that captures impression management. The net effect of image usage on sophisticated shareholder proposals is therefore positive and stronger for specific images than for generic images. This evidence is *not* consistent with successful impression management. We explore several possible explanations for institutional activists increasing activism in relation to images, including confounding factors such as the activists 1) targeting low-quality-reporting firms with proposals to increase the firms' disclosures (Flammer et al., 2021) and 2) treating ESG-focused and non-ESG-focused shareholder proposals differently. However, our inferences (untabulated) remain the same when we remove the eight ESG-reporting proposals, as when we focus solely on the 49 ESG-activity proposals.

After ruling out these topic-based explanations, we conjecture that the channel of influence is informational, since information often increases divergence in perceptions of ESG performance (Christensen et al., 2022). In a direct effect, shareholder activists could use the information in specific images to predict future negative performance, a relationship that we documented in Table 7. In a spillover effect, shareholder activists could use information about one ESG dimension to infer deficiency in other ESG dimensions. For example, they could infer lack of employee diversity from specific images of employee volunteer days. The informational channel for aesthetic images remains conjectural because AI is limited in its ability to link objects to specific ESG dimensions, much less to extrapolate to other dimensions. This is another intriguing avenue for future research. Together, the shareholder proposal results suggest that unsophisticated individual activists fall prey to visual impression management while sophisticated institutional activists extract information from images.[8]

5.3. External accolades

Sponsors of ESG awards typically invite judges with experience and credentials in ESG-relevant fields, including high-ranking investment analysts and managers, ESG-advisory consultants and lawyers, nonprofit leaders, sustainability advocates, and business school academics (IJGlobal, 2023; Cohn Reznick, 2023; ESG Investing: Sustainability News Events & Awards, 2024; ESG & Sustainability Awards, 2024). We conservatively expect that some of the judges read ESG reports. Importantly, this setting allows us to examine the perceptions of a wide range of individuals and civil society organizations that fall outside the

[6] A well-developed literature concludes that individual investors are less sophisticated than institutions (e.g., Bartov et al., 2000; Bushee, 1998; Hand, 1990; Walther, 1997). On ESG topics in particular, though, evidence is still emerging. Surveys conclude that most institutional investors consider ESG information in their decisions (Amel-Zadeh & Serafeim, 2018). For retail traders, archival evidence is mixed and incomplete: Robinhood investors do not respond to ESG disclosures at all (Moss, Naughton, and Wang, 2024), but general retail investors respond to financially material ESG news (Li et al., 2023; Serafeim & Yoon, 2022). There is little systematic evidence on where investors obtain their ESG information (e.g., what percentage of investors of each type rely on in-house research vs. third-party ratings, to what extent investors rely on automated aggregation or manual processing). These are important questions for future survey research.

[7] We tabulate the sample proposals by sponsor type in Table OA5 Panel A of the Online Appendix. The incentives and motives of religious organizations and labor unions are complex, but these groups are not necessarily unsophisticated—they may have financial sophistication in addition to social objectives. As a result, these groups defy easy characterization and ex-ante predictions. Ex post, we do not observe any significant coefficients on image-related variables for these groups, as we report in Table OA5 Panel B.

[8] We widen the shareholder activism setting to a more general population, all mutual funds, in Table OA6 of the Online Appendix. The results on mutual fund voting on shareholder proposals are inconclusive; we do not find that mutual fund support for management varies with ESG report image usage. However, we cannot disentangle the attribution of this result. Various contributing factors include the following: 1) mutual funds that vote on proposals are a different shareholder population from funds that raise proposals, and therefore have different priorities and concerns; 2) mutual funds may not always take the time to read individual ESG reports; and 3) the mutual funds that do read ESG reports may see through visual impression management.

investor community (Cerrato & Ferrando, 2020; European Commission, 2013).

In Table 7, we replace the dependent variable in Model 3 with ESG_AWARD, an indicator variable equal to 1 if the firm wins one or more ESG awards in the year after report publication. We find that ESG awards are associated with better ESG performance and more extensive textual disclosure (as evidenced in the significantly negative coefficient on POOR_ESG and the significantly positive coefficient on TEXT_CONTENT, respectively). The significantly positive coefficient on TEXT_SPECIFICITY provides direct evidence that award givers read ESG reports.

In Columns 1–2, we find that image usage increases a firm's chances of winning ESG awards in the face of negative ESG news; this effect is driven by specific images (Columns 3–4). The favorable perception associated with image usage could reflect ESG award judges giving weight to substantiating information or falling susceptible to impression management. In contrast, Column 5 shows that heavy generic image usage is associated with lower chances of winning ESG awards, regardless of ESG performance. For an average firm with 67.15 as its SCORE_CONTROVERSIES, a one-standard-deviation increase in specific image usage is associated with a 46.28 %–49.85 % higher chance of winning at least one ESG award. Regardless of ESG performance, a one-standard-deviation increase in generic IMAGES/PAGE is associated with a 73.31 % lower chance of winning at least one ESG award. Together, these results suggest that ESG award committees value specific images and penalize generic images.

In summary, we conclude that visual impression management in ESG reports enhances perception in retail traders but rarely achieves the same in institutional shareholders. A small number of experts—institutional shareholder activists and ESG award committees—even react negatively, which suggests that corporations and researchers may have underestimated the detection costs of visual impression management.

6. Conclusion

Using state-of-the-art deep learning and AI image analysis, we systematically document the determinants and consequences of ESG report image usage across a large sample of firms from 2005 to 2018.

ESG reports that display lower-quality textual content tend to use more images. Firms with stronger motivations for impression management, such as those belonging to problematic industries, issuing non-GRI reports, or demonstrating poor ESG performance, tend to use more aesthetic images (especially generic images, which contain little substance). In combination, these findings add nuance to inferences from prior literature on determinants of ESG disclosure. Even as firms operating in polluting industries and sin industries (Gamerschlag et al., 2011; Grougiou et al., 2016) and firms experiencing negative events like oil spills and nuclear catastrophes (Bonetti et al., 2023; Heflin & Wallace, 2017; Patten, 1992) increase their textual and numerical transparency, they could also be increasing attempts at visual impression management.

Heavier image usage also is associated with future ESG problems that are not yet incorporated into current performance assessments. Moreover, the actual number of negative ESG performance incidents could be greater, as our measure captures extreme, negative ESG performance that results in financial penalties. These findings consistently confirm the concern, raised in experimental and small-scale studies, that impression management is an important motivation for image usage in corporate reports. At the same time, our findings that reports use high percentages of specific images, which provide substantiating information for decision-making, temper the concern that image usage is equivalent to impression management.

Lastly, we find evidence that image usage affects stakeholder decisions through two mechanisms. Image usage affects unsophisticated audiences like retail investors and individual shareholder activists through impression management, thereby improving their perception of ESG performance. In contrast, image usage affects financially sophisticated and ESG-prioritizing experts predominantly through informational channels. Institutional shareholder activists use specific images to inform their proposals. ESG award committees react positively to informative images but negatively to strategic images, which is consistent with their seeing through visual impression management—perhaps even perceiving it as analogous to narrative or textual embellishment. However, activists constitute only a small fraction of institutional shareholders, and most institutional shareholders do not appear to react to images. Whether they remain unaffected because they carefully attend to and see through visual impression management or because they pay little attention to the ESG report and its presentation choices is inconclusive.

A natural question arises: Why do corporations persist in ESG report visual impression management when its effectiveness is so uncertain? The greenwashing literature shares this puzzle with the earnings management and "cheap talk" literatures (e.g., Farrell and Rabin (1996), Arya et al. (1998), Dewatripoint and Tirole (2005), Lo (2008), and Sobel (2015)). Like Bothello et al. (2023), we find evidence that at least some stakeholders fail to see through impression management some of the time. Since the direct and indirect costs of visual impression management are low, even small benefits can outweigh the costs. In exchange for these potential benefits, firms appear to be quite willing to risk engaging in impression management. Simultaneously, managers may fail to accurately predict the costs of visual impression management, especially if they are unaware that it can elicit negative reactions.

Our tests of stakeholder perceptions have some limitations. First, measures of shareholder responses are noisier in the ESG report setting than in the annual report setting, as we are unable to conduct short-window event studies. We do not know the exact dates of ESG report publication, only the month and year. Second, empirically testing experimental insights necessitates an imperfect translation of settings. For example, a direct translation of an experimental setting to an empirical one would involve testing consumer reactions, but consumers generally do not read ESG reports (Leonelli et al., 2023). Third, our settings for testing consequences do not perfectly overlap, so we cannot consistently compare results across settings. For example, the institutional investors who trade shares in the returns setting and the institutional investors who engage in activism in the proposal setting are interrelated, yet different, populations. We cannot expect uniform results. Fourth, in the tests where we do not observe coefficients that are significantly different from zero, we cannot rule out empirical confounds such as sample composition, model specifications, and lack of statistical power. A natural consequence of testing a large, generalizable sample is that we observe an average effect; non-average effects may emerge or dominate in narrower samples. Despite this caveat, we believe model misspecification and lack of power are unlikely to explain our results, since our models stem from substantial prior literature and our sample is not small. Fifth, ESG reports are composed of mixed and complex topics; there is room to further disentangle stakeholder reactions to separate pieces of bundled information. Lastly, we acknowledge the limitations of ESG ratings as imperfect measures of ESG performance (Berg et al., 2022; Chatterji et al., 2009; Christensen et al., 2022; Raghunandan & Rajgopal, 2023; Serafeim and Yoon, 2023). We address this concern by including two external-consequence-based measures (media controversies and regulatory penalties) in our main tests and an alternative ESG rating in a robustness test. Nevertheless, the question of how well ESG ratings capture ESG activities and achievements remains a major concern for researchers and practitioners alike.

More broadly, visual impression management falls under the manipulation of presentation in the taxonomy of impression management (Merkl-Davies & Brennan, 2007). Studies on the forms of impression management have rarely examined more than one at a time. While these behaviors share similar motivations, we lack theory and evidence on their relationships, e.g., as complements or substitutes. Such interplays present rich opportunities for future research. Relatedly, survey evidence on the thought and work processes of corporate report producers would enrich our understanding of report inputs. Are report producers aware of the benefits and costs to specific or generic images? How do they balance or prioritize different stakeholder audiences? Do they actively make choices to court awards?

We hope that our approach to combining automated and manual methods informs future research, with the caveat that these measures would benefit from more refinement. Like other researchers using big data methods (e.g., Adukia et al., 2023), we face a trade-off in data collection: automated methods provide consistency across researchers and over time, while qualitative and small-scale studies capture more nuances. Also, our ability to draw conclusions on the independent roles of specific and generic images is limited because the high empirical correlation between them necessitates relative and incremental interpretations.

We invite AI researchers to train models that recognize specific images and generic images, and to replicate our manual classification on a larger scale. We look forward to future research on image characteristics that were not examined in this paper, such as their placement, their reinforcement of each other, and their alignment with the surrounding text. We believe that extracting information content from images and examining negative reactions to visual impression management sketch substantial directions for revising prior beliefs. Most recently, the application of generative AI to image creation represents a new frontier in visual research. While companies now find it increasingly easy to produce realistic-looking specific images using AI prompts, audiences will find it harder to identify images with information content. Whether audiences can discern specific images reflecting real-world events from ostensibly specific images created by AI, or whether the proliferation of AI image creation cheapens specific images entirely, will be a big question in visual impression management.

Declaration of interest

IRB letters are not applicable. No human subjects were studied in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aos.2025.101616.

Appendix I. Examples of ESG Report Images in Collected Data

Below are examples of 5 broad categories of images commonly found in ESG reports: (1) images related to people, (2) images related to products, (3) images related to ESG initiatives, (4) marginally relevant decorative images, and (5) data visualizations.

We refer to Types (1)–(4), images without numerical data, collectively as "aesthetic images". In this paper, we focus on the information dissemination and impression management roles of aesthetic images.

Exhibits (1)-(3) also serve as examples of specific images, while Exhibit (4) serves as an example of a generic image.

Data visualizations and aesthetic images are created and processed differently (Avgerinou & Petterson, 2011). For example, in data visualizations, numerical or relational information such as scale, direction, and area are primary; in aesthetic images, retinal information such as hue, saturation, and texture are primary (Carter, Hipwell, and Quinnell, 2012). The main purpose of aesthetic images is to "evoke feelings and attitudes" (Carter, Hipwell, and Quinnell, 2012). The processing of data visualizations requires numeracy, the ability to understand numbers, while the processing of aesthetic images does not.

Exhibit (1). Images Related to People
Top Left: CEO; Bottom Left: Global Sustainability Director; Right: Employee
Source: Whirlpool Corporation 2017 Sustainability Report

Exhibit (2). Images Related to Products
Top Left: Product Exterior; Bottom Left: Product Interior; Bottom Right: Product Award
Source: Whirlpool Corporation 2017 Sustainability Report

Exhibit (3). Images Related to Initiatives
Top Left: Windmills; Top Right: Energy-Efficient Headquarters Building; Bottom Left: Corporate Philanthropy; Bottom Right: Employee Volunteer Activities
Source: Whirlpool Corporation 2017 Sustainability Report

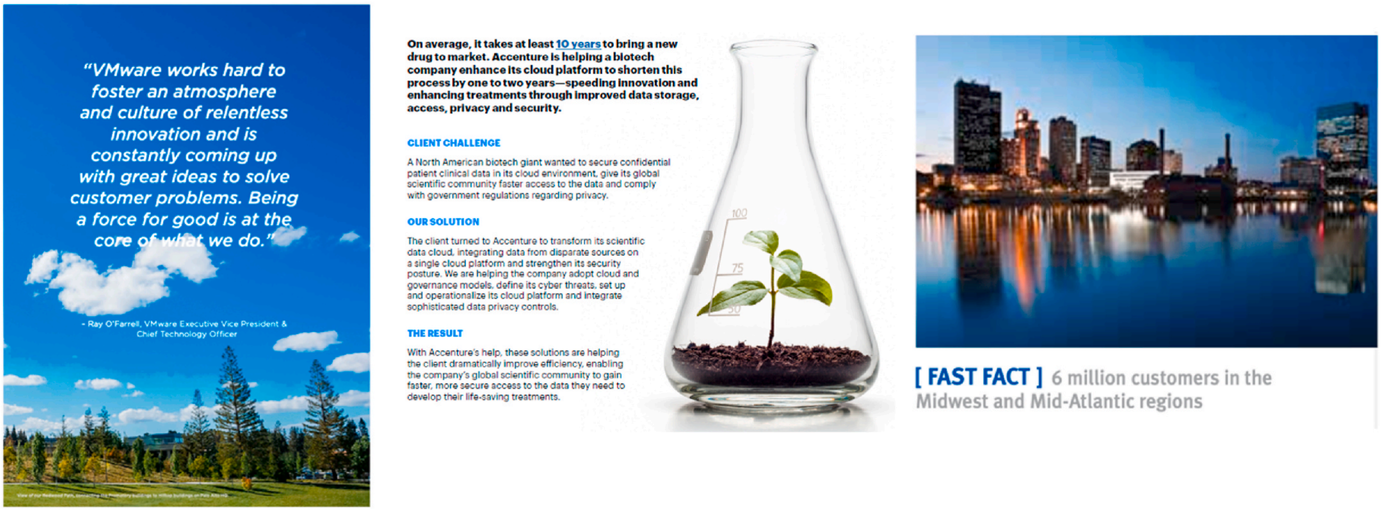

Exhibit (4). Marginally Relevant Decorative Images
Above Left: Group of trees and blue sky; Above Middle: Beaker with small plant growing out of layer of soil inside; Above Right: Skyline of city
Sources: VMWare Global Impact Report 2015; Accenture Corporate Citizenship Report 2017; FirstEnergy Sustainability Report 2012

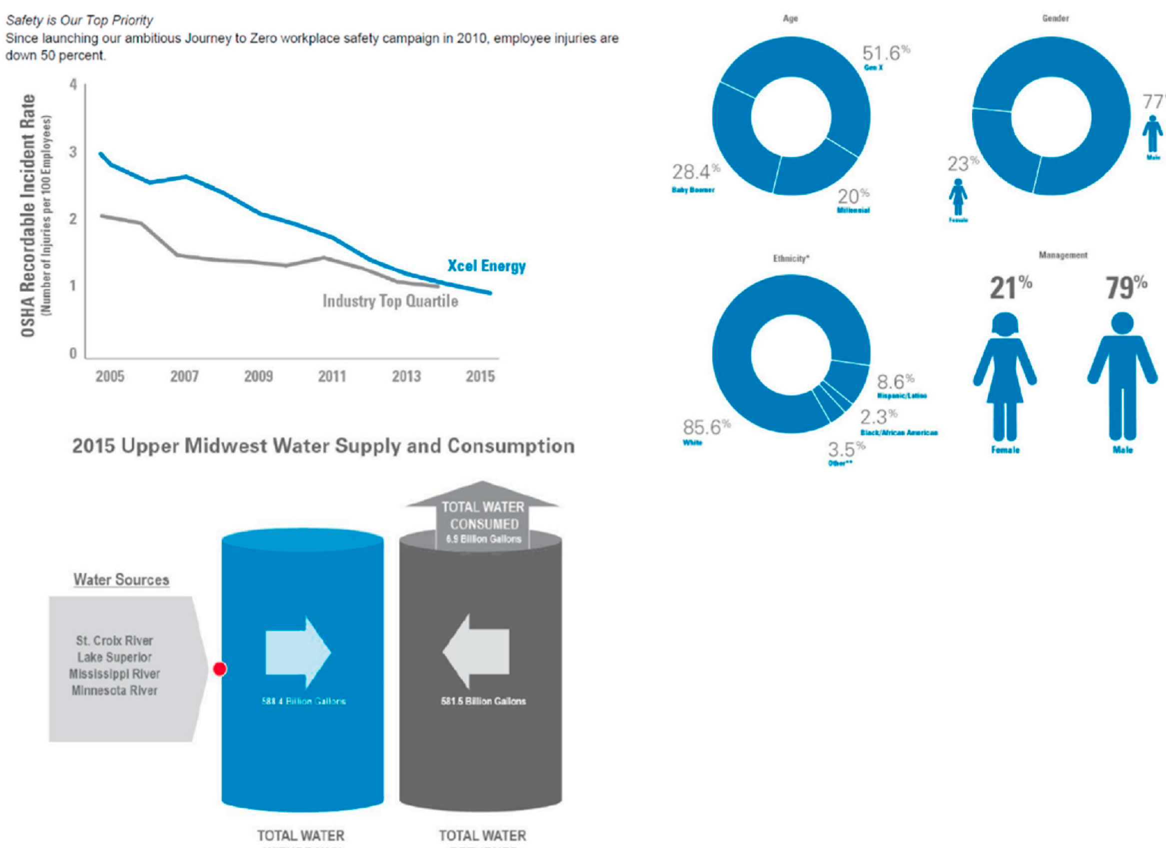

Exhibit (5). Images Related to Data Visualization
Top Left: Graph Depicting Workplace Injury Trend Compared to Industry Top Quartile; Top Right: Charts Depicting Personnel Characteristic Distributions; Bottom: Diagram Depicting Water Supply Consumption in One Operating Region
Source: Xcel Energy Corporation 2015 Corporate Responsibility Report

References for* *Appendix I

Avgerinou, M. D., and R. Petterson. (2011). Towards a cohesive theory of visual literacy. *Journal of Visual Literacy*, 30(2): 1–19.

Carter, L., P. Hipwell, Patricia, and L. Quinnell. (2012). A picture is worth a thousand words: An approach to learning about visuals. *Australian Journal of Middle Schooling*, 12(2): 4–15.

Appendix II. Automated Image Identification Method and Manual Refinement

1. Steps

1.1. Recognition by File Type

First, we convert PDF reports into Microsoft Word documents using two different types of conversion software. In rare cases when conversion fails with both types of software, we eliminate the reports from the sample. Next, we convert Word files into ZIP folders using Python. Using Python, we delete images below 20 KB in size, which are mostly design elements such as lines, logos, and banners. We manually scrolled through extracted images in 100 randomly selected reports and found that they rarely contain data visualizations such as tables, charts, or graphs, which are often stored in text format. Previous drafts of the paper were based on this basic method of extracting and counting images. To our knowledge, we were the first to apply this method to business research in 2020. However, since the inception of this paper, algorithms for recognizing images have evolved rapidly with the advancement of artificial intelligence (AI). Beginning in the September 2023 revision of the paper, we replace previous measures of image usage with refined measures detailed below.

1.2. Refinement by Object Detection

We use one of the state-of-the-art deep learning models of object detection, the Faster R–CNN + InceptionResNetV2 network, to refine the recognition of aesthetic images. Faster R–CNN is trained on the Open Images V4 dataset introduced by Kuznetsova et al. (2020), a dataset of 9.2 M Flickr images that the authors annotated and validated. Classifications are available for 600 categories of objects. Each detected object is listed with its detected category (e.g., human) up to 100 objects, and a detection score ranging from 0 to 1 indicating the confidence of the detection. Faster R–CNN has been extensively validated in the computer science field and is recognized as one of the top performers in object detection (Huang et al., 2017; Han et al., 2018). It has also been demonstrated in applications ranging from detecting pedestrians (Shao et al., 2021), to detecting fabric defects (Zhou et al., 2020), to detecting skin cancer (Hartanto and Wibowo, 2020), among others.

We input the 133,431 images extracted earlier using Python into the Faster R–CNN network, over a runtime of 3 months. We accept detections with confidence thresholds over 0.5, i.e., the object in the image *more likely than not* fits the detected category. We associate each image with and only with the object detection category reporting the highest detection confidence, subject to the floor of 0.5. Thus, using object detection, we identify 81,982 images (61.44 % of the Python-identified sample) as aesthetic images.

We manually verify that small illustrations of objects, which are commonly used in data visualizations, are not detected as objects. Large illustrations of people and landscapes are detected as containing objects. We accept these detections, given that we would classify them as aesthetic images manually as well. However, some issues with false detection remain, such as graphs being detected as objects. Next, we sample manually by category for the dual purposes of 1) identifying high-inaccuracy categories and 2) quantifying the information properties of aesthetic images.

1.3. Refinement by Manual Sampling

FasterR-CNN labels objects in images according to a taxonomy, proceeding from the highest level of detail to the highest levels of aggregation, as illustrated in Figure S1 below (Open Images, 2018). For example, an image of a football will receive the labels of "football" (object level), "ball" (Level 2, alongside cricket balls, volleyballs, etc.), and "sports equipment" (Level 1).

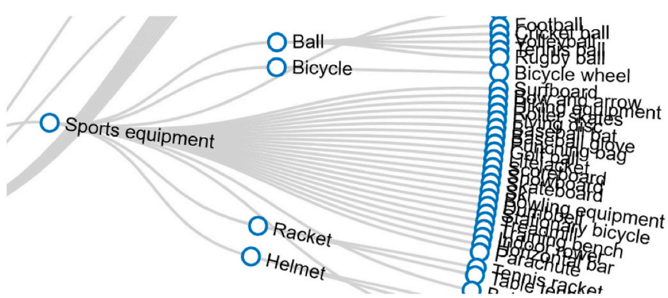

Fig. S1. *Magnified Illustration of the Object Classification Hierarchy (Google, 2018)* https://storage.googleapis.com/openimages/2018_04/bbox_labels_600_hierarchy_visualizer/circle.html

We construct a plan for manual sampling based on the distribution of detected categories in sample images. As Figure S2 shows, 68 % of sample images belong to the top 8 detection categories with more than 2000 images each. The remaining 32 % of images is split among more than 50 detection categories. For categories with fewer than 200 images, we aggregate them to their Level 2 label. If a Level 2 detection category does not reach 200 images, we aggregate again to Level 1. If a Level 1 category still does not reach 200 images, we drop that category. Thus, we delete 3531 images that cannot be aggregated. For the resulting 78,451 sample images split across 62 categories, we randomly select 1 % per category or 20 images per category, whichever is greater, for manual sampling.

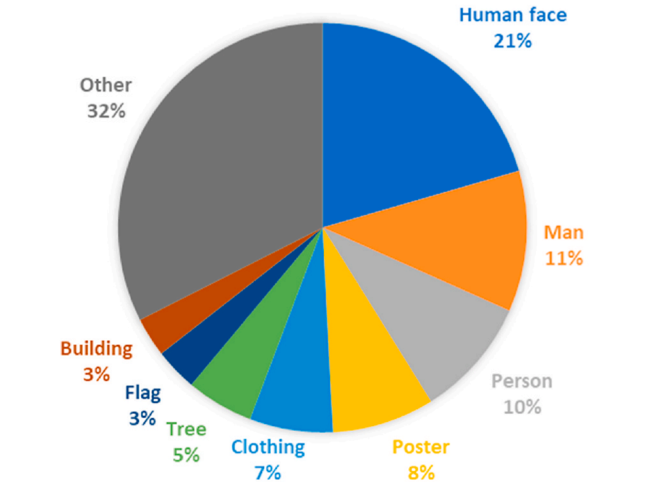

Fig. S2. Distribution of Object Identified Images by Category

Each of the 3 co-authors independently looks through the 1610 images randomly selected for manual sampling. We reach consensus to delete 8 highly inaccurate categories of 12,143 images in which almost none of the manually sampled images are identified correctly. For example, we found most images labeled as Flag, Ball, Window, or Plate to be tables or charts, and most images labeled as Candle to be bar graphs. Below, Table S1 details the frequency of each category in the sample, and the number of images manually sampled in each category. Highly inaccurate categories are marked as "Yes" under the column "Deleted".

Table S1
Manual Sampling by Computer Identified Category

Category	Level of Classification	Computer Identified Frequency	Percentage of Sample	Manual Sampling Frequency	Deleted	Specificity Weight
Human face	Object	16,105	20.53 %	161		79.92 %
Man	Object	8737	11.14 %	87		79.69 %
Person	Object	7420	9.46 %	74		77.48 %
Poster	Object	6328	8.07 %	63	Yes	–
Clothing	Object	5165	6.58 %	52		76.28 %
Tree	Object	4165	5.31 %	42		30.95 %
Flag	Object	2652	3.38 %	27	Yes	–
Building	Object	2428	3.09 %	24		91.67 %
Woman	Object	1940	2.47 %	20		85.00 %
Car	Object	1329	1.69 %	20		83.33 %
Plant	Object	1172	1.49 %	20		40.00 %
Footwear	Object	998	1.27 %	20		78.33 %
Suit	Object	928	1.18 %	20		91.67 %
Land vehicle	Object	928	1.18 %	20		78.33 %
Airplane	Object	743	0.95 %	20		50.00 %
Skyscraper	Object	720	0.92 %	20		33.33 %
Flower	Object	598	0.76 %	20		23.33 %
Toy	Object	596	0.76 %	20		53.33 %
Ball	Object	595	0.76 %	20	Yes	–
Window	Object	582	0.74 %	20	Yes	–
House	Object	577	0.74 %	20		51.67 %
Truck	Object	546	0.70 %	20		93.33 %
Mobile phone	Object	519	0.66 %	20	Yes	–
Boat	Object	496	0.63 %	20		81.67 %
Tower	Object	456	0.58 %	20		66.67 %
Plate	Object	451	0.57 %	20	Yes	–
Traffic sign	Object	447	0.57 %	20	Yes	–
Vehicle	Object	407	0.52 %	20		63.33 %
Bottle	Object	399	0.51 %	20		83.33 %
Train	Object	369	0.47 %	20		86.67 %
Helmet	Object	362	0.46 %	20		98.33 %
Wheel	Object	361	0.46 %	20		58.33 %
Balloon	Object	351	0.45 %	20		15.00 %
Bird	Object	350	0.45 %	20		13.33 %
Chair	Object	331	0.42 %	20		82.50 %
Whiteboard	Object	307	0.39 %	20	Yes	–
Candle	Object	262	0.33 %	20	Yes	–
Food	Object	251	0.32 %	20		60.00 %
Table	Object	251	0.32 %	20		85.00 %
Boy	Object	246	0.31 %	20		58.33 %
Laptop	Object	240	0.31 %	20		28.33 %

(continued on next page)

Table S1 (*continued*)

Category	Level of Classification	Computer Identified Frequency	Percentage of Sample	Manual Sampling Frequency	Deleted	Specificity Weight
Book	Object	237	0.30 %	20		46.67 %
Jeans	Object	236	0.30 %	20		81.67 %
Box	Object	234	0.30 %	20		53.33 %
Picture frame	Object	234	0.30 %	20		13.33 %
Snack	Object	224	0.29 %	20		95.00 %
Fashion accessory	Object	222	0.28 %	20		51.67 %
Bicycle	Object	212	0.27 %	20		53.33 %
Human hand	Object	204	0.26 %	20		21.67 %
Fashion accessory	L2	573	0.73 %	20		51.67 %
Mammal	L2	558	0.71 %	20		28.33 %
Container	L2	351	0.45 %	20		78.33 %
Computer Equipment	Manually combined from 3 L1 categories	256	0.33 %	20		60.00 %
Vehicle-General	L1	466	0.59 %	20		80.00 %
Food-General	L1	365	0.47 %	20		43.33 %
Animal-General	L1	324	0.41 %	20		0.00 %
Office supplies	L1	296	0.38 %	20		23.33 %
Furniture	L1	294	0.37 %	20		60.00 %
Plant	L1	292	0.37 %	20		40.00 %
Clothing-General	L1	285	0.36 %	20		78.33 %
Auto part	L1	242	0.31 %	20		50.00 %
Tool-General	L1	238	0.30 %	20		37.50 %
Total		**78,451**	**100.00 %**	**1610**	**n/a**	**n/a**

1.4. Estimating Specific vs. Generic Image Usage

During manual sampling, we also independently record what percentage of images appear to be firm-specific according to professional judgment. For example, aesthetic images that substantiate ESG claims can convey information—photos of the firm's recycled products and energy-efficient infrastructure substantiate the firm's environmental commitment, and photos of the firm's employees volunteering or of community members benefitting from philanthropy substantiate the firm's social initiatives. Such images can contain impression management elements as well; duly enhancing perception through substantiating information and unduly enhancing perception through evoking positive feelings cannot be disentangled in specific images. However, we contrast specific images against generic images, stock images or generic illustrations that evoke ESG themes without concrete information on whether the firm has undertaken such activities. Generic images are likely to contain more visual impression management elements than specific images.

We consider images to be specific if the detected objects sport brands and names, or otherwise appear to be personalized to the firm, as opposed to appearing to be a stock image. When we encounter large divergence of specificity estimates, we discuss our approaches to seek consensus. If we reach consensus, we take the average of all three co-authors' specificity percentages. If we cannot reach consensus, we take the two majority opinions and average those two percentages. The resulting specificity weights for each category are detailed in Table S1.

To estimate the number of specific images, we weight the total number of object-detected images by the specificity percentage of each image's category. Conversely, we calculate Generality Weight = 100 % – Specificity Weight and apply the generality weight of each category to estimate generic image usage.

Note that specific image usage and generic image usage are correlated by construction but not multicollinear. Theoretically, if we hold total image usage constant, specific image usage and generic image usage should be negatively correlated. However, as we discuss in the paper, we observe empirically that specific image usage and generic image usage are positively correlated. In other words, reports that use more specific images tend to use more generic images too, resulting in higher image usage overall.

Table S2
Correlations between Types of Images

		(1)	(2)	(3)	(4)	(5)	(6)
(1)	IMAGES_PER_PAGE (all)	1.000					
(2)	IMAGES_PER_PAGE (specific)	0.916*	1.000				
(3)	IMAGES_PER_PAGE (generic)	0.856*	0.794*	1.000			
(4)	IMAGES_TO_TEXT (all)	0.798*	0.724*	0.685*	1.000		
(5)	IMAGES_TO_TEXT (specific)	0.752*	0.821*	0.656*	0.926*	1.000	
(6)	IMAGES_TO_TEXT (generic)	0.687*	0.635*	0.811*	0.883*	0.837*	1.000

1.5. Out-of-Sample Validation

We perform out-of-sample validation for the specificity-weighting method. First, we sort the categories by frequency, to focus on the categories that have the largest impact on our sample. Then, within the most frequent categories, we identify two extremes: the three object detection categories with the highest specificity weights (human face - 79.92 %, man – 79.69 %, and person – 77.48 %) and the three categories with the lowest specificity scores (plant – 40 %, skyscraper – 33.33 %, and tree – 30.95 %). From each extreme, we randomly select 30 images that are outside the previous manually examined sample. We then examine these out-of-sample images using the classification steps detailed in Section 1.3. For each inconclusive image, we locate the ESG report from which it was extracted and read the image captions for context. We find that 67 % of the images in the higher-specificity group are specific, compared to 20 % in the lower-specificity group. Note that the benchmark for the higher-specificity (lower-specificity) group is not 100 % (0 %), as less than 100 % (more than 0 %) in those groups are classified as specific in Table S1. This out-of-sample examination increases our confidence in the validity of the sampling-based specificity measure and our empirical strategy to exploit relative specificity across categories.

2. Discussion

2.1. Relation to Image Measures in Previous Drafts

We document extremely high correlation between our image measures as refined by deep learning object detection and as extracted by the previous Python method. The correlation is 0.840 for IMAGES_PER_PAGE and 0.867 for IMAGES_TO_TEXT. Moreover, our results and inferences remain similar. Thus, researchers who do not have the time or computing resources to process object detection may use Python image-counting as a quick estimate. Object detection has allowed us to take our original analysis farther by estimating the information content of aesthetic images using a combination of AI and manual methods.

2.2. Exploration of Heuristics for Image Specificity

While we encourage readers to apply the specificity weights in Table S1 to analyze ESG reports, we recognize that many may not have the time or resources to engage in extensive object detection. In this section, we report on descriptive findings relating specificity to other image characteristics that may be of wide interest.

2.3. Image Size

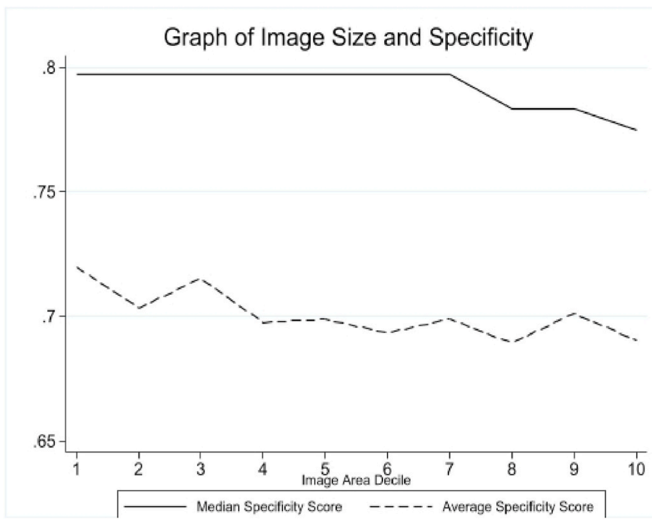

Fig. S3. Graph of Image Size and Specificity

Table S3
Regression of Image Size on Specificity Weight

VARIABLES	(1)
	SPECIFICITY WEIGHT
IMAGE SIZE	−0.0034***
	(0.0009)
Constant	0.7054***
	(0.0012)
Observations	66,243
Adjusted R-squared	0.0755
Year FE	YES
Report FE	YES
Firm FE	YES

While manually sampling images, we noticed that specific images are often snapshot-sized and embedded in text, while some generic images had been enlarged for aesthetic purposes. We explore whether a pattern exists by graphing image size against specificity in Figure S3 and regressing image size on specificity in Table S3.

We focus on the size of images as they appear to readers, due to its relevance to visual impression management, rather than the digital size of images in pixels or bytes. We use Python to calculate image size as the width multiplied by the length of the image in inches, i.e., the area of the image, and sort image size into deciles to reduce noise. In Figure S3, we observe that the median image specificity weight is stable for the lower (<8) deciles of image size, but decreases sharply from the 8th to the top deciles. The average image specificity weight decreases steadily from the lowest to the highest size deciles.

In Table S3, we run an OLS regression with specificity weight as the dependent variable and image size as the independent variable of interest. Note that this regression is at the image level, not at the report level or firm level. To control for report and firm characteristics, we add related fixed effects and year fixed effects. Standard errors are clustered by report. We find a strongly significant negative coefficient on image size, indicating that it is inversely related to specificity.

Overall, it appears that images with high information content are on average smaller; the largest images are most likely to be generic.

2.4. Industry Representation in the Top and Bottom Image Specificity Quartiles

We undertake another descriptive analysis by exploring industry representation at the highest and lowest extremes of image specificity. First, we calculate percentage representations of each Fama-French industry in the entire sample in Column (1). Then, we separate ESG reports into the top and bottom quartile in specificity weights (see Table S1). We calculate each industry's percentage representation in the top quartile of report image specificity in Column (2), and in the bottom quartile of report image specificity in Column (3). We highlight any industry where representation in either extreme quartile deviates by 200 basis points or more from representation in the overall sample. We format higher-than-expected specificity (i.e., overrepresentation in the high-specificity quartile or underrepresentation in the low-specificity quartile) with italics and underlining. Lower-than-expected specificity (i.e., underrepresentation in the high-specificity quartile or overrepresentation in the low-specificity quartile) is bolded and shaded in gray.

Table S4
Industry Concentration in the Top and Bottom Image Specificity Quartiles

Industry		(1)	(2)	(3)
		All %	High Specificity %	Low Specificity %
1	Business Equipment	14 %	15 %	**18 %**
2	Chemicals and Allied Products	7 %	**5 %**	**10 %**
3	Consumer Durables	3 %	4 %	*<u>1 %</u>*
4	Consumer Non-Durables	8 %	**4 %**	8 %
5	Finance	14 %	*<u>18 %</u>*	13 %
6	Healthcare, Medical Equipment, and Drugs	7 %	*<u>13 %</u>*	6 %
7	Manufacturing	11 %	**9 %**	12 %
8	Oil, Gas, and Coal Extraction and Products	5 %	4 %	*<u>1 %</u>*
9	Telecommunications	2 %	3 %	2 %
10	Utilities	9 %	**5 %**	8 %
11	Wholesale, Retail, and Some Services	9 %	9 %	10 %
12	Other	12 %	11 %	12 %
	Total	100 %	100 %	100 %

A notable theme from this table is that industries with potentially heavy pollution lag in image specificity: business equipment, chemicals and allied products, consumer non-durables, manufacturing, and utilities are underrepresented in the top specificity quartile or overrepresented in the lowest specificity quartile. This is broadly consistent with our finding in Table 3 and Section 4.2 of the paper that polluting industries tend to use more generic images. Surprisingly, consumer durables, finance, healthcare, and oil, gas, and coal lead in ESG report image specificity; they are over-represented in the top specificity quartile or underrepresented in the lowest specificity quartile. Upon further reflection, ESG reports in the finance industry often feature photos of employees or volunteering activities, while those in oil, gas, and coal often feature photos of trucks, helmets, and other equipment. Another intriguing observation is that consumer durables lead in image specificity, and yet consumer non-durables lag in image specificity. This distinction would suggest that B2C industries are not a monolith, and durable and non-durable producers approach ESG report images differently.

To summarize, with the exception of consumer durables and oil, gas, and coal, heavy manufacturing industries tend to lag in ESG report image specificity. Service industries like finance and healthcare lead in ESG report image specificity, though they showcase people instead of objects.

References for Appendix II. Google. 2018. Hierarchy visualizer. https://storage.googleapis.com/openimages/2018_04/bbox_labels_600_hierarchy_visualizer/circle.html. Accessed June 1, 2023.

Han, J., D. Zhang, G. Cheng, N. Liu, and D. Xu. 2018. Advanced deep-learning techniques for salient and category-specific object detection: A survey. *IEEE Signal Processing Magazine*, 35(1):84–100.

Hartanto, C.A. and A. Wibowo. 2020. Development of mobile skin cancer detection using Faster R–CNN and MobileNet v2 Model. *International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*.

Huang, J., V. Rathod, C. Sun et al., 2017. Speed/accuracy trade-offs for modern convolutional object detectors. *Conference on Computer Vision and Pattern Recognition (CFPR)*.

Kuznetsova, A., H. Rom, N. Alldrin et al. 2020. The Open Images Dataset V4. *International Journal of Computer Vision,* 128:1956–1981.

Shao, X., J. Wei, D. Guo et al., 2021. Pedestrian detection algorithm based on improved Faster RCNN. *IEEE 5*th *Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*.

Zhou, H., B. Jang, Y. Chen, and D. Troendle. 2020. Exploring Faster RCNN for Fabric Defect Detection. *Third International Conference on Artificial Intelligence for Industries (AI4I)*.

Appendix III. Variable Definitions

Variable	Definition
Image-Related Variables	
(1) IMAGES	Number of unique images no smaller than 20 KB in a ESG report. Equals 0 if no images fitting the criteria are detected.
(2) PAGES	Number of pages in a ESG report.
(3) WORDS (in thousands)	Number of words in a ESG report divided by 1000.

(*continued on next page*)

(*continued*)

Variable	Definition
(4) IMAGES/PAGE	Number of unique images no smaller than 20 KB in a ESG report scaled by the number of pages. IMAGES/PAGE (specific) and IMAGES/PAGE (generic) denote IMAGES/PAGE weighted by the specificity or generality of the images' object detection categories, respectively. Please see Appendix II for classification and estimation methodology.
(5) IMAGES/TEXT	Number of unique images no smaller than 20 KB in a ESG report scaled by the number of words in thousands. IMAGES/TEXT (specific) and IMAGES/TEXT (generic) denote IMAGES/TEXT weighted by the specificity or generality of the images' object detection categories, respectively. Please see Appendix II for classification and estimation methodology.
(6) AR_IMAGES/PAGE	Number of unique images no smaller than 20 KB in an annual report scaled by the number of pages. Equals 0 if no images fitting the criteria are detected.
(7) AR_IMAGES/TEXT	Number of unique images no smaller than 20 KB in an annual report scaled by the number of words, further divided by 1000. Equals 0 if no images fitting the criteria are detected.
ESG Report Characteristics	
(8) GRI	Equals 1 if the ESG report declares it follows Global Reporting Initiative standards, and 0 otherwise.
(9) TEXT_CONTENT	ESG Disclosure Score from Bloomberg, created by counting the number of topics disclosed in each ESG report from a predetermined list of important topics and standardizing the count to a scale of 0–100.
(10) TEXT_SPECIFICITY	A report's SPECIFICITY percentile minus its BOILERPLATE percentile each year. We measure BOILERPLATE as the number of words in boilerplate sentences divided by the total number of words in each report. We define boilerplate language as four-word phrases that are commonly used among ESG reports. We classify a phrase as a boilerplate phrase if it appears in more than 10 % of the ESG reports in a specific year. We manually review and exclude common phrases such as "as a result of", "as well as the", "Global Reporting Initiative GRI", etc., from boilerplate phrases. We classify a sentence that contains one or more boilerplate phrases and does not contain any numbers as boilerplate sentences. We measure SPECIFICITY as the number of specific words, mostly numerals and proper nouns, divided by the total number of words in each report. We identify specific words using the Stanford Name Entity Recognizer.
ESG Performance Measures	
(11) SCORE_ESG	Refinitiv ESG Score for fiscal year t.
(12) SCORE_CONTROVERSIES	100 – monthly Refinitiv ESG Controversies Score averaged over fiscal year t.
(13) MISCONDUCT	Log (1 + total incidences of corporate misconduct aggregated in the database Violation Tracker, which places emphasis on ESG issues resulting in regulatory infractions, government fines, and class action lawsuits).
(14) POOR_ESG	100 – SCORE_ESG, i.e., 100 minus the Refinitiv ESG Score.
Industry Characteristics	
(15) B2C_IND	Equals 1 if the company operates in the consumer goods industries (SIC 0000–0999, 2000–2399, 2500–2599, 2700–2799, 2830–2869, 3000–3219, 3420–3429, 3523, 3600–3669, 3700–3719, 3751, 3850–3879, 3880–3999, 4813, 4830–4899, 5000–5079, 5090–5099, 5130–5159, 5220–5999, 7000–7299, 7400–9999) or finance industries (6000–6999), and 0 otherwise.
(16) CONTROVERSIAL_IND	Equals 1 if the company operates in the weapons (SIC 3760–3769, 3795, 3480–3489), oil (1300, 1310–1339, 1370–1382, 1389, 2900–2912, 2990–2999), cement (3240–3241), or biotech (2833–2836) industries, and 0 otherwise.
(17) DIRTY_IND	Equals 1 if the company operates in the manufacturing, mining, or chemicals industries (SIC 2000–3999), and 0 otherwise.
(18) HIGHDIFF_IND	Equals 1 if the company is in an industry with higher than median product differentiation among all firms in the data provided by Hoberg and Phillips (2016) that year, and 0 otherwise.
(19) SIN_IND	Equals 1 if the company operates in the alcohol (SIC 2100–2199), tobacco (2080–2085), or gambling (NAICS 7132, 71312, 713120, 71329, 713290, 72112, 721120) industries, and 0 otherwise.
Firm Characteristic Controls	
(20) ADVT	Advertising expense/beginning total assets. Set to 0 if missing in Compustat.
(21) AGE	Log (1 + number of years the firm has been in the CRSP database).
(22) ANALYSTS	Average number of analysts following the firm during the fiscal year. Set to 0 if no analyst following is found in I/B/E/S.
(23) BTM	Book value of equity/market value of equity.
(24) CAPX	Capital expenditures/total assets beginning balance.
(25) INSTHOLDINGS	Percentage of shares owned by institutional investors.
(26) LEV	(Long term debt + debt in current liabilities)/beginning total assets.
(27) MKTCAP	Log (1 + market capitalization).
(28) NPR	Number of firm press releases during the fiscal year.
(29) RD	R&D expenditures/beginning total assets. Set to 0 if missing in Compustat.
(30) ROA	Income before extraordinary items/beginning of year total assets.
(31) SALES	Net sales/beginning of year total assets.
(32) SALESGROWTH	$Sales_t/Sales_{t\text{-}1} - 1$.
(33) TURNOVER	Average daily trading turnover in fiscal year t, where daily turnover = (vol/(1000*shrout)) * 100.
Stakeholder Perception Outcomes	
(34) CAR	$CAR = \sum_{t=1}^{12} R_t - ER_t$ over the 12 months following ESG report publication date, where expected returns are estimated using the Fama-French Three Factor Model.
(35) ESG_AWARD	Indicator variable equal to 1 if the company receives an ESG award, recorded by Refinitiv, and 0 otherwise.
(36) PROPOSALS	Log (1 + total number of shareholder proposals).

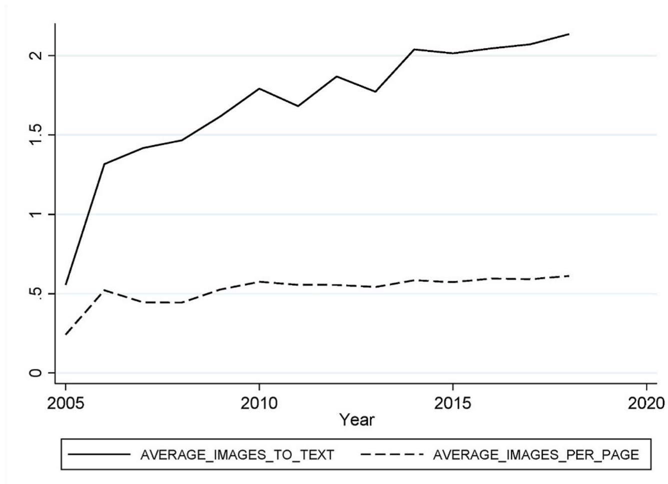

Fig. 1. 2005–2018 Average Image Usage in Sample ESG Reports

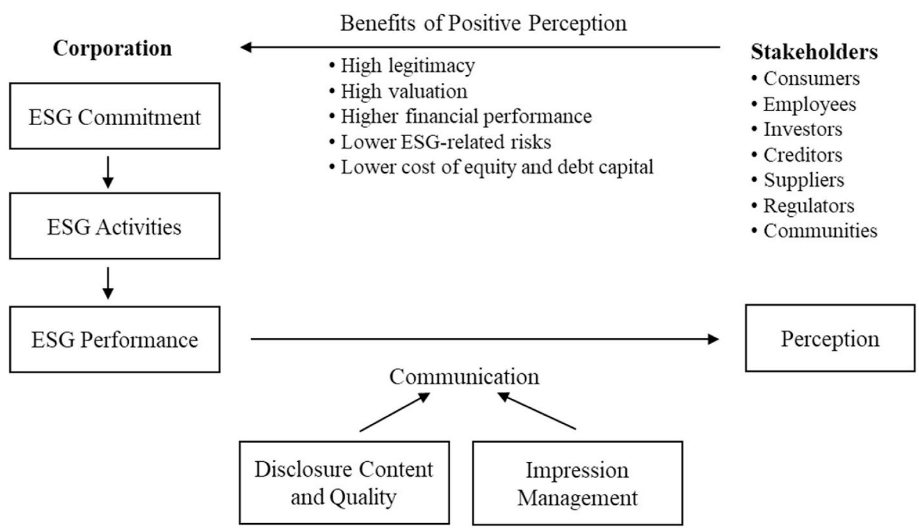

Fig. 2. Conceptual Diagram: Impression Management Enhances Stakeholder Perception of Corporate ESG Performance

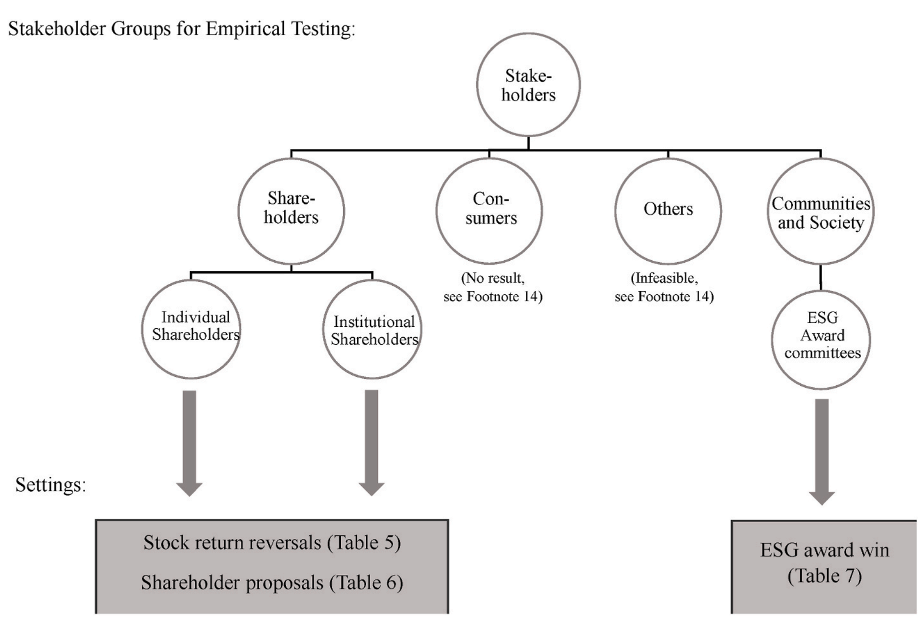

Fig. 3. Stakeholder Groups, Settings, and Tests

Table 2
ESG Report Image Usage and Textual Characteristics
This table reports estimation of the OLS regression *Image Usage*$_{i,t}$ = α + β_1*TEXT_CONTENT*$_{i,t}$ + β_2*TEXT_SPECIFICITY*$_{i,t}$ + β_3*AR_Image Usage* + β_4*MKTCAP*$_{i,t}$ + β_5*AGE*$_{i,t}$ + β_6*BTM*$_{i,t}$ + β_7*LEV*$_{i,t}$ + β_8*ADVT*$_{i,t}$ + β_9*RD*$_{i,t}$ + β_{10}*CAPX*$_{i,t}$ + β_{11}*ROA*$_{i,t}$ + β_{12}*SALES*$_{i,t}$ + β_{13}*SALESGROWTH*$_{i,t}$ + β_{14}*TURNOVER*$_{i,t}$ + β_{15}*INSTHOLDINGS*$_{i,t}$ + β_{16}*ANALYSTS*$_{i,t}$ + β_{17}*NPR*$_{i,t}$ + Industry Fixed Effects + Year Fixed Effects. *Image Usage*$_{i,t}$ is scaled by number of pages in Columns 1, 3, 5 (*IMAGES/PAGE*$_{i,t}$) and scaled by units of thousand-words in Columns 2, 4, 6 (*IMAGES/TEXT*$_{i,t}$).
Measures of image usage contain all sample object-detected images in Columns 1 & 2, focus on image usage weighted by specificity in Columns 3 & 4, and focus on image usage weighted by generality in Columns 5 & 6. Moreover, regressions on specific image usage (Columns 3 & 4) control for generic image usage, and regressions on generic image usage (Columns 5 & 6) control for specific image usage. Coefficient standard errors, presented in parentheses, are clustered by firm. Significance at the 0.01, 0.05 and 0.1 levels in two-tailed tests is denoted by ***, **, and *, respectively.

Image Type	All		Specific		Generic	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	IMAGES/PAGE	IMAGES/TEXT	IMAGES/PAGE	IMAGES/TEXT	IMAGES/PAGE	IMAGES/TEXT
TEXT_CONTENT	−0.0083***	−0.0402***	−0.0013***	−0.0063***	−0.0004**	−0.0014**
	(0.0009)	(0.0036)	(0.0004)	(0.0014)	(0.0002)	(0.0006)
TEXT_SPECIFICITY	0.0005**	0.0035***	0.0000	0.0005	0.0000	0.0002
	(0.0002)	(0.0008)	(0.0001)	(0.0003)	(0.0000)	(0.0002)
AR_IMAGES/PAGE	0.0052		0.0181		−0.0083	
	(0.0331)		(0.0120)		(0.0051)	
AR_IMAGES/TEXT		−0.0310		0.0123		−0.0103*
		(0.0315)		(0.0127)		(0.0056)
IMAGES/PAGE (generic)			1.7501***			
			(0.0556)			
IMAGES/TEXT (generic)				1.7358***		
				(0.0593)		
IMAGES/PAGE (specific)					0.3587***	
					(0.0073)	
IMAGES/TEXT (specific)						0.3872***
						(0.0090)
MKTCAP	−0.0033	0.0025	0.0114**	0.0504***	−0.0059***	−0.0245***
	(0.0123)	(0.0471)	(0.0049)	(0.0177)	(0.0022)	(0.0081)
AGE	0.0050	−0.0298	−0.0049	−0.0339*	0.0031	0.0137
	(0.0132)	(0.0542)	(0.0054)	(0.0197)	(0.0023)	(0.0089)
BTM	0.0105	0.1496	−0.0090	0.0060	0.0058	0.0137
	(0.0410)	(0.1531)	(0.0178)	(0.0649)	(0.0088)	(0.0323)
LEV	0.1139**	0.5239**	−0.0153	0.0067	0.0223*	0.0543
	(0.0568)	(0.2324)	(0.0262)	(0.0942)	(0.0120)	(0.0436)
ADVT	−0.2057	0.5871	0.1430	0.9093	−0.1002	−0.3945
	(0.5596)	(2.4705)	(0.2431)	(0.8544)	(0.0888)	(0.3674)
RD	−0.0841	1.2123	−0.4056**	−1.3735**	0.1900***	0.8339***
	(0.3611)	(1.5741)	(0.1583)	(0.5362)	(0.0682)	(0.2775)
CAPX	0.4329*	2.7502**	0.0892	0.7175	0.0123	−0.0402
	(0.2583)	(1.2295)	(0.1212)	(0.4748)	(0.0510)	(0.1947)
ROA	0.1181	0.0203	−0.0095	−0.0668	0.0184	0.0284
	(0.1706)	(0.7658)	(0.0687)	(0.2413)	(0.0302)	(0.1188)
SALES	0.0213	0.1586	0.0245**	0.0846**	−0.0094**	−0.0254
	(0.0204)	(0.0963)	(0.0096)	(0.0345)	(0.0043)	(0.0176)
SALESGROWTH	0.0080	0.1976	−0.0241	−0.0540	0.0122	0.0484
	(0.0496)	(0.2093)	(0.0193)	(0.0761)	(0.0092)	(0.0351)
TURNOVER	−0.0138	−0.0883	−0.0021	0.0093	−0.0009	−0.0150
	(0.0155)	(0.0618)	(0.0069)	(0.0249)	(0.0032)	(0.0110)
INSTHOLDINGS	0.0636	0.3595**	−0.0115	−0.0016	0.0133**	0.0396*
	(0.0392)	(0.1433)	(0.0147)	(0.0487)	(0.0065)	(0.0232)
ANALYSTS	−0.0007	−0.0026	0.0004	0.0010	−0.0003	−0.0008
	(0.0018)	(0.0067)	(0.0007)	(0.0026)	(0.0003)	(0.0013)
NPR	0.0038	0.0086	0.0005	−0.0004	0.0003	0.0012
	(0.0029)	(0.0111)	(0.0010)	(0.0037)	(0.0005)	(0.0017)
Constant	0.7387***	2.2986***	0.0049	−0.1136	0.0916***	0.3153***
	(0.1360)	(0.5203)	(0.0516)	(0.1848)	(0.0224)	(0.0838)
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	2602	2602	2602	2602	2602	2602
Adjusted R-squared	0.1138	0.1661	0.6663	0.7235	0.6690	0.7211

Table 1
Descriptive Statistics
This table presents the sample construction of aesthetic images and statistical distributions of main variables. Please see Appendix I for variable definitions and Appendix II for details on the object detection and image classification procedures. Panel A summarizes how we gathered the images in the study.
Panel B presents the sample distribution by Fama-French 12-Industry Classifications. Panel C presents summary statistics. All continuous variables are winsorized at 1 % and 99 %, except for INSTHOLDINGS, which is winsorized to a maximum of 1 following recommendations by Wharton Research Data Services (WRDS, 2020). Panel D presents an abbreviated correlation table with pairwise correlations for key characteristics of sample ESG reports and sample firms. All variables are constructed at the firm-year level. Significance at the 0.1 level is denoted by *.

Panel A. Sample Construction of Aesthetic Images

Images extracted from 2602 sample reports using Python	133,431
Less: No objects detected with confidence >0.5	(51,449)
Less: Highly inaccurate object detection as determined by manual sampling (across 8 categories)	(12,143)
Total object-detected images	**69,839**
Less: Category frequency too scattered for meaningful aggregation	(3531)
Images with specificity and generality weights (54 categories)	**66,308**

Panel B. Sample Distribution by Fama-French 12-Industry Classification

			Mean					
	Image Type		All		Specific		Generic	
			(1)	(2)	(3)	(4)	(5)	(6)
	Industry	N	IMAGES/PAGE	IMAGES/TEXT	IMAGES/PAGE	IMAGES/TEXT	IMAGES/PAGE	IMAGES/TEXT
(1)	Consumer Non-Durables	214	0.67	2.18	0.43	1.41	0.20	0.65
(2)	Consumer Durables	70	0.59	1.97	0.40	1.31	0.16	0.54
(3)	Manufacturing	284	0.58	1.99	0.38	1.29	0.17	0.59
(4)	Oil, Gas, Coal Extraction & Products	142	0.59	1.80	0.41	1.24	0.16	0.50
(5)	Chemicals & Allied Products	176	0.64	2.16	0.40	1.35	0.19	0.65
(6)	Business Equipment	355	0.46	1.61	0.30	1.04	0.13	0.47
(7)	Telephone and Television Transmission	51	0.63	2.35	0.44	1.64	0.17	0.61
(8)	Utilities	224	0.55	1.69	0.35	1.09	0.17	0.51
(9)	Wholesale, Retail, & Some Services	223	0.58	2.10	0.38	1.38	0.17	0.62
(10)	Healthcare, Medical Equipment & Drugs	188	0.58	1.88	0.40	1.30	0.15	0.50
(11)	Finance	354	0.55	1.93	0.36	1.27	0.16	0.57
(12)	Other	321	0.57	1.85	0.38	1.23	0.17	0.54
	Total	**2602**	**1.03**	**3.46**	**0.37**	**1.26**	**0.16**	**0.56**

Panel C. Summary Statistics

		N	Mean	Standard Deviation	Min	P25	Median	P75	Max
	Image-Related Variables								
(1)	IMAGES (raw)	2602	26.66	21.03	0.00	11.00	22.00	36.00	108.00
(2)	PAGES	2602	55.38	39.78	4.00	27.00	45.00	74.00	201.00
(3)	WORDS (in thousands)	2602	18.70	15.03	1.24	7.77	14.78	25.02	77.77
(4a)	IMAGES/PAGE (all)	2602	0.57	0.35	0.00	0.32	0.53	0.77	1.75
(4b)	IMAGES/PAGE (specific)	2602	0.37	0.24	0.00	0.20	0.35	0.51	1.17
(4c)	IMAGES/PAGE (generic)	2602	0.16	0.11	0.00	0.09	0.15	0.22	0.55
(5a)	IMAGES/TEXT (all)	2602	1.91	1.42	0.00	0.92	1.61	2.56	7.24
(5b)	IMAGES/TEXT (specific)	2602	1.26	0.96	0.00	0.57	1.05	1.69	4.75
(5c)	IMAGES/TEXT (generic)	2602	0.56	0.45	0.00	0.25	0.47	0.72	2.51
(6)	AR_IMAGES/PAGE	2602	0.19	0.28	0.00	0.03	0.08	0.20	1.39
(7)	AR_IMAGES/TEXT	2602	0.48	0.95	0.00	0.05	0.15	0.37	5.31
	ESG Report Characteristics								
(8)	GRI	2602	0.66	0.47	0.00	0.00	1.00	1.00	1.00
(9)	TEXT_CONTENT	2602	39.34	12.50	14.05	30.17	39.67	47.93	68.86
(10)	TEXT_SPECIFICTY	2602	0.01	46.17	−91.00	−35.00	0.00	35.00	93.00
	ESG Performance Measures								
(11)	SCORE_ESG	2166	69.57	12.42	33.31	61.59	70.81	78.63	92.21
(12)	SCORE_CONTROVERSIES	2166	67.15	25.50	31.25	41.86	75.00	92.86	99.66
(13)	MISCONDUCT	2111	0.88	0.93	0.00	0.00	0.69	1.39	3.85
	Firm Characteristic Controls								
(14)	MKTCAP	2602	9.65	1.38	5.97	8.76	9.63	10.57	12.64
(15)	AGE	2602	3.45	0.79	0.81	2.98	3.65	4.00	4.51
(16)	BTM	2602	0.44	0.32	−0.20	0.22	0.37	0.59	1.62
(17)	LEV	2602	0.30	0.18	0.00	0.17	0.28	0.40	0.94
(18)	ADVT	2602	0.01	0.02	0.00	0.00	0.00	0.01	0.11
(19)	RD	2602	0.02	0.04	0.00	0.00	0.00	0.02	0.19
(20)	CAPX	2602	0.04	0.04	0.00	0.02	0.04	0.06	0.20
(21)	ROA	2602	0.06	0.07	−0.18	0.02	0.05	0.09	0.24

(*continued on next page*)

Table 1 (*continued*)

Panel C. Summary Statistics

		N	Mean	Standard Deviation	Min	P25	Median	P75	Max
(22)	SALES	2602	0.87	0.71	0.04	0.38	0.71	1.15	3.70
(23)	SALESGROWTH	2602	0.04	0.15	−0.39	−0.03	0.03	0.10	0.68
(24)	TURNOVER	2602	0.98	0.66	0.26	0.57	0.78	1.15	4.08
(25)	INSTHOLDINGS	2602	0.71	0.26	0.00	0.65	0.78	0.87	1.00
(26)	ANALYSTS	2602	17.12	7.87	1.33	11.83	16.83	22.00	38.75
(27)	NPR	2602	14.12	4.11	5.00	12.00	14.00	16.00	29.00
	Stakeholder Perception Outcomes								
(28)	CAR	1327	−0.01	0.24	−1.54	−0.12	0.00	0.12	1.09
(29)	PROPOSALS	2111	0.74	0.73	0.00	0.00	0.69	1.10	2.56
(30)	ESG_AWARD	1860	0.80	0.40	0.00	1.00	1.00	1.00	1.00

Panel D. Abbreviated Correlations between Key Variables

	Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1)	IMAGES/PAGE	1.00												
(2)	IMAGES/TEXT	0.81*	1.00											
(3)	AR_IMAGES/PAGE	0.01	−0.02	1.00										
(4)	AR_IMAGES/TEXT	0.00	−0.02	0.91*	1.00									
(5)	TEXT_CONTENT	−0.25*	−0.32*	−0.02	−0.04*	1.00								
(6)	TEXT_SPECIFICITY	0.03	0.07*	−0.03	−0.03	0.07*	1.00							
(7)	MKTCAP	−0.10*	−0.12*	0.15*	0.09*	0.36*	0.07*	1.00						
(8)	SCORE_ESG	−0.21*	−0.26*	0.03	0.00	0.46*	0.04*	0.46*	1.00					
(9)	SCORE_CONTROVERSIES	−0.09*	−0.11*	0.04*	0.02	0.17*	0.10*	0.44*	0.27*	1.00				
(10)	MISCONDUCT	0.06*	0.01	0.08*	0.03	0.10*	0.07*	0.26*	0.12*	0.25*	1.00			
(11)	CAR	0.03	0.01	0.02	0.01	0.02	0.02	0.03	0.03	0.04	0.00	1.00		
(12)	PROPOSALS	−0.07*	−0.11*	0.10*	0.05*	0.20*	0.05*	0.62*	0.31*	0.43*	0.38*	0.04	1.00	
(13)	ESG_AWARDS	−0.07*	−0.13*	0.05*	0.03	0.26*	0.11*	0.22*	0.30*	0.19*	0.16*	0.10*	0.18*	1.00

Table 3

ESG Report Image Usage and Group Tendencies for Impression Management

This table reports results from OLS regressions to test differences in ESG report image usage between groups with differing motivations or opportunities for impression management. Below, Columns 1–10 report estimation of the OLS regression *Image Usage*$_{i,t}$ = α + β1*Industry Group*$_{i,t}$ + β2*TEXT_CONTENT*$_{i,t}$ + β3*TEXT_SPECIFICITY*$_{i,t}$ + *Firm Controls*$_{i,t}$ + Year Fixed Effects, where *Industry Group*$_{i,t}$ = *SIN_IND*$_{i,t}$, *DIRTY_IND*$_{i,t}$, *CONTROVERSIAL_IND*$_{i,t}$, *B2C_IND*$_{i,t}$, or *COMPETITIVE_IND*$_{i,t}$. Fama-French 12-industry fixed effects are excluded.

SIN_IND is an indicator variable equal to 1 if the firm operates in the alcohol, tobacco, or gambling industry and 0 otherwise. *DIRTY_IND* is an indicator variable equal to 1 if the firm operates in the manufacturing, mining, or chemicals industry and 0 otherwise. *CONTROVERSIAL_IND* is an indicator variable equal to 1 if the firm operates in the weapons, oil, cement, or biotech industry and 0 otherwise. *B2C_IND* is an indicator variable equal to 1 if the firm operates in the consumer goods or finance industry and 0 otherwise. *HIGHDIFF_IND* is an indicator variable equal to 1 if the firm's product differentiation is above the median among all firms within each fiscal year in the data provided by Hoberg and Phillips (2016), and 0 otherwise.

Columns 11 & 12 report estimation of the regression *Image Usage*$_{i,t}$ = α + β1*GRI*$_{i,t}$ + β2*TEXT_CONTENT*$_{i,t}$ + β3*TEXT_SPECIFICITY*$_{i,t}$ + *Firm Controls*$_{i,t}$ + Industry Fixed Effects + Year Fixed Effects. *GRI* is an indicator variable equal to 1 if the ESG report follows GRI Sustainability Reporting Standards, and 0 otherwise.

Throughout this table, image usage measures contain all object-detected images. Univariate comparisons between groups and re-estimation of these regression models on specific and generic images are available in Tables OA1 and OA2 of the Online Appendix. Coefficient standard errors, presented in parentheses, are clustered by firm. Significance at the 0.01, 0.05 and 0.1 levels in two-tailed tests is denoted by ***, **, and *, respectively.

Type of Images	All											
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT
SIN_IND	0.0908**	0.3282*										
	(0.0444)	(0.1847)										
DIRTY_IND			0.0860***	0.2669***								
			(0.0244)	(0.0980)								
CONTROVERSIAL_ IND					0.0729**	0.1875						
					(0.0317)	(0.1233)						
B2C							0.0350	0.1643*				
							(0.0250)	(0.0981)				
HIGHDIFF_IND									0.0354*	0.1772**		
									(0.0207)	(0.0773)		
GRI											−0.0912***	−0.4271***
											(0.0232)	(0.0943)
TEXT_CONTENT	−0.0079***	−0.0395***	−0.0083***	−0.0407***	−0.0080***	−0.0399***	−0.0077***	−0.0388***	−0.0083***	−0.0415***	−0.0064***	−0.0311***
	(0.0009)	(0.0036)	(0.0009)	(0.0035)	(0.0009)	(0.0035)	(0.0009)	(0.0036)	(0.0009)	(0.0036)	(0.0010)	(0.0040)

(*continued on next page*)

Table 3 (*continued*)

Type of Images	All											
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT	IMAGES/ PAGE	IMAGES/ TEXT
TEXT_SPECIFICITY	0.0004**	0.0032***	0.0005**	0.0034***	0.0004**	0.0033***	0.0004**	0.0033***	0.0004**	0.0034***	0.0004**	0.0032***
	(0.0002)	(0.0008)	(0.0002)	(0.0008)	(0.0002)	(0.0008)	(0.0002)	(0.0008)	(0.0002)	(0.0008)	(0.0002)	(0.0008)
Firm Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	–	–	–	–	–	–	–	–	–	–	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2602	2602	2602	2602	2602	2602	2602	2602	2531	2531	2602	2602
Adjusted R-squared	0.0985	0.1569	0.1075	0.1619	0.1003	0.1571	0.0988	0.1581	0.1006	0.1601	0.1244	0.1804

Table 4

ESG Report Image Usage and Performance Motivations for Impression Management

This table explores the relationship between ESG report image usage and ESG performance. Panel A examines the relationship between concurrent image usage and ESG performance in year t. Panel B examines whether ESG report image usage in t has predictive power for ESG performance in t+1.

In Panel A, we explore whether the relationship between image usage and SCORE_ESG varies monotonically. We assign a value of 1 to BOTTOM_QUARTILE_SCORE_ESG, SECOND_QUARTILE_SCORE_ESG, THIRD_QUARTILE_SCORE_ESG if the observation's SCORE_ESG falls within the 1st-25th, 26-50th, or 51-75th percentile in each year, respectively. Otherwise, we set these indicator variables to 0. We regress these 3 indicator variables on ESG report image usage, along with firm controls, Fama-French 12-industry fixed effects, and year fixed effects. Observations in the top quartile of SCORE_ESG are omitted.For conciseness, *IMAGE_USAGE* alternately takes on the value of IMAGES/PAGE in columns labeled "Images Scaled By … Pages" or the value of IMAGES/TEXT in columns labeled "Images Scaled by … Text". Columns 1 & 2 include all object-detected image usage in the dependent variable. Columns 3 & 4 focus on specific image usage as the dependent variable while controlling for generic image usage. Columns 5 & 6 focus on generic image usage as the dependent variable while controlling for specific image usage. Standard errors, presented in parentheses, are clustered at the firm level. Significance at the 0.01, 0.05 and 0.1 levels in two-tailed tests is denoted by ***, **, and *, respectively.

Panel A. Monotonical Relationship between SCORE_ESG and Image Usage

Image Type	All		Specific		Generic	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	IMAGES/PAGE	IMAGES/TEXT	IMAGES/PAGE	IMAGES/TEXT	IMAGES/PAGE	IMAGES/TEXT
BOTTOM_QUARTILE_SCORE_ESG	0.0882***	0.4701***	−0.0015	0.0266	0.0114*	0.0406*
	(0.0298)	(0.1151)	(0.0134)	(0.0452)	(0.0059)	(0.0224)
SECOND_QUARTILE_SCORE_ESG	0.0625**	0.2839***	0.0095	0.0425	0.0032	0.0129
	(0.0276)	(0.1036)	(0.0106)	(0.0370)	(0.0044)	(0.0163)
THIRD_QUARTILE_SCORE_ESG	0.0628***	0.3180***	0.0094	0.0496	0.0032	0.0131
	(0.0229)	(0.0872)	(0.0093)	(0.0326)	(0.0038)	(0.0140)
SCORE_CONTROVERSIES	0.0001	−0.0003	0.0000	−0.0002	0.0000	0.0000
	(0.0004)	(0.0014)	(0.0002)	(0.0006)	(0.0001)	(0.0003)
TEXT_CONTENT	−0.0085***	−0.0424***	−0.0019***	−0.0089***	−0.0002	−0.0005
	(0.0011)	(0.0039)	(0.0004)	(0.0016)	(0.0002)	(0.0007)
TEXT_SPECIFICITY	0.0006***	0.0042***	0.0001	0.0008**	0.0000	0.0001
	(0.0002)	(0.0009)	(0.0001)	(0.0004)	(0.0000)	(0.0002)
Control for Generic Image Usage	–	–	YES	YES	–	–
Control for Specific Image Usage	–	–	–	–	YES	YES
Firm Controls	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	2166	2166	2166	2166	2166	2166
Adjusted R-squared	0.1395	0.2074	0.6814	0.7299	0.6794	0.7205

In Panel B, we report estimation of the model $MISCONDUCT_{i,t+1} = \alpha + \beta_1 Image\ Usage_{i,t} + \beta_2 Image\ Usage_{i,t} * POOR\_ESG_{i,t} + \beta_3 Image\ Usage_{i,t} * SCORE\_CONTROVERSIES_{i,t} + \beta_4 POOR\_ESG_{i,t} + \beta_5 SCORE\_CONTROVERSIES_{i,t} + \beta_6 TEXT\_CONTENT_{i,t} + \beta_7 TEXT\_SPECIFICITY_{I,t} + Firm\ Controls_{i,t} + Industry\ Fixed\ Effects_i + Year\ Fixed\ Effects_t + \varepsilon$. MISCONDUCT refers to the natural log of incidents of fines, penalties, and lawsuit settlements due to corporate misconduct against stakeholders recorded by the database Violation Tracker. POOR_ESG is 100 – SCORE_ESG, i.e., 100 minus the Refinitiv ESG Score.

Panel B. Predictive Power of ESG Report Image Usage for Future Negative ESG Events

Image type	All		Specific		Generic	
Scaled by	Pages	Text	Pages	Text	Pages	Text
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	MISCONDUCT					
IMAGE_USAGE	0.5805***	0.1198**	0.7247**	0.1546*	1.4315**	0.2638
	(0.2068)	(0.0533)	(0.3316)	(0.0889)	(0.7223)	(0.1776)
*IMAGE_USAGE**POOR_ESG	−0.0088**	−0.0019*	−0.0131**	−0.0027*	−0.0221*	−0.0039
	(0.0041)	(0.0011)	(0.0059)	(0.0015)	(0.0132)	(0.0033)

(*continued on next page*)

Table 4 (*continued*)

In Panel B, we report estimation of the model $MISCONDUCT_{i,t+1} = \alpha + \beta_1 Image\ Usage_{i,t} + \beta_2 Image\ Usage_{i,t} * POOR\_ESG_{i,t} + \beta_3 Image\ Usage_{i,t} * SCORE\_CONTROVERSIES_{i,t} + \beta_4 POOR\_ESG_{i,t} + \beta_5 SCORE\_CONTROVERSIES_{i,t} + \beta_6 TEXT\_CONTENT_{i,t} + \beta_7 TEXT\_SPECIFICITY_{I,t} + Firm\ Controls_{i,t} + Industry\ Fixed\ Effects_i + Year\ Fixed\ Effects_t + \varepsilon$. MISCONDUCT refers to the natural log of incidents of fines, penalties, and lawsuit settlements due to corporate misconduct against stakeholders recorded by the database Violation Tracker. POOR_ESG is 100 – SCORE_ESG, i.e., 100 minus the Refinitiv ESG Score.

Panel B. Predictive Power of ESG Report Image Usage for Future Negative ESG Events

Image type	All		Specific		Generic	
Scaled by	Pages	Text	Pages	Text	Pages	Text
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	MISCONDUCT					
*IMAGE_USAGE**SCORE_CONTROVERSIES	−0.0019	−0.0003	−0.0022	−0.0005	−0.0055	−0.0011
	(0.0020)	(0.0005)	(0.0028)	(0.0007)	(0.0068)	(0.0016)
POOR_ESG	0.0039	0.0026	0.0038	0.0023	0.0025	0.0011
	(0.0036)	(0.0033)	(0.0035)	(0.0033)	(0.0034)	(0.0032)
SCORE_CONTROVERSIES	0.0038***	0.0034**	0.0036***	0.0033**	0.0036***	0.0033***
	(0.0014)	(0.0013)	(0.0014)	(0.0013)	(0.0013)	(0.0013)
TEXT_CONTENT	0.0019	0.0018	0.0021	0.0020	0.0020	0.0019
	(0.0027)	(0.0028)	(0.0027)	(0.0028)	(0.0027)	(0.0028)
TEXT_SPECIFICITY	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Control for Generic Image Usage	–	–	YES	YES	–	–
Control for Specific Image Usage	–	–	–	–	YES	YES
FIRM CONTROLS	YES	YES	YES	YES	YES	YES
IND FE	YES	YES	YES	YES	YES	YES
YEAR FE	YES	YES	YES	YES	YES	YES
Observations	2111	2111	2111	2111	2111	2111
Adjusted R-squared	0.3994	0.3972	0.3998	0.3973	0.3991	0.3966

Table 5

ESG Report Image Usage and Shareholder Perception – Return Reversals

This table reports results from estimating OLS regressions using the following model: $CAR_{i,t+1} = \alpha + \beta_1 Image\ Usage_{i,t} + \beta_2 Image\ Usage_{i,t} * POOR\_ESG_{i,t} + \beta_3 Image\ Usage_{i,t} * SCORE\_CONTROVERSIES_{i,t} + \beta_4 POOR\_ESG_{i,t} + \beta_5 SCORE\_CONTROVERSIES_{i,t} + \beta_6 TEXT\_CONTENT_{i,t} + \beta_7 TEXT\_SPECIFICITY_{i,t} + Firm\ Controls_{i,t} + Industry\ Fixed\ Effects_i + Year\ Fixed\ Effects_t + \varepsilon$. We calculate CAR as the excess of realized returns over Fama-French Three-Factor expected returns (estimated as $RF_t + \beta MktRF_t + s(SMB_t) + h(HML_t)$, where RF_t is the one-month T-bill return, $MktRF_t$ is the excess value-weighted market return from the WRDS Fama-French database, SMB_t is the return on factor-mimicking portfolio for firm size, and HML_t is the return on factor-mimicking portfolio for book-to-market). β is estimated using monthly stock returns in the 36 months before fiscal year end, requiring at least six observations for estimation. We include only December 31 year-end firms, constituting 73.60 % of all sample observations.

We partition the sample by yearly median institutional ownership to reflect investor sophistication. For conciseness, *IMAGE_USAGE* alternately takes on the value of IMAGES/PAGE in columns labeled "Images Scaled By … Pages" or the value of IMAGES/TEXT in columns labeled "Images Scaled by … Text". Columns with the header "Image Type … All" include all object-detected images in measures of image usage. Columns with the header "Image Type … Specific" focus on specific image usage while controlling for generic image usage. Columns with the header "Image Type … Generic" focus on generic image usage while controlling for specific image usage. Significance at the 0.01, 0.05 and 0.1 levels in two-tailed tests is denoted by ***, **, and *, respectively.

Subsample	High Institutional Ownership						Low Institutional Ownership					
Image Type	All		Specific		Generic		All		Specific		Generic	
Images Scaled By	Pages	Text	Pages	Text	Pages	Text	Pages	Text	Pages	Text	Pages	Text
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR
IMAGE_USAGE	0.1172	0.0316	0.2343	0.0618	0.0247	0.0019	0.3471***	0.0617*	0.4858***	0.0933*	1.0147**	0.1276
	(0.1372)	(0.0386)	(0.2038)	(0.0601)	(0.4816)	(0.1244)	(0.1192)	(0.0357)	(0.1797)	(0.0542)	(0.4262)	(0.1157)
*IMAGE_USAGE**POOR_ESG	−0.0021	−0.0008	−0.0046	−0.0015	−0.0023	−0.0010	−0.0040**	−0.0001	−0.0054*	0.0000	−0.0107*	−0.0000
	(0.0028)	(0.0007)	(0.0039)	(0.0010)	(0.0095)	(0.0023)	(0.0020)	(0.0006)	(0.0028)	(0.0009)	(0.0061)	(0.0015)
*IMAGE_USAGE**SCORE_CONTROVERSIES	−0.0000	0.0001	−0.0007	0.0001	0.0024	0.0007	−0.0025**	−0.0006**	−0.0035**	−0.0009**	−0.0085***	−0.0020**
	(0.0014)	(0.0004)	(0.0019)	(0.0006)	(0.0047)	(0.0012)	(0.0010)	(0.0003)	(0.0015)	(0.0005)	(0.0032)	(0.0008)
POOR_ESG	0.0015	0.0020	0.0020	0.0023	0.0007	0.0009	0.0021	−0.0003	0.0018	−0.0006	0.0016	−0.0004
	(0.0015)	(0.0015)	(0.0014)	(0.0014)	(0.0015)	(0.0014)	(0.0015)	(0.0013)	(0.0014)	(0.0014)	(0.0015)	(0.0012)
SCORE_CONTROVERSIES	0.0008	0.0007	0.0011	0.0007	0.0004	0.0004	0.0013*	0.0011	0.0012	0.0011	0.0013*	0.0010
	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0007)	(0.0007)
TEXT_CONTENT	0.0028***	0.0027***	0.0028***	0.0029***	0.0027***	0.0027***	−0.0001	0.0000	−0.0000	0.0001	−0.0001	0.0001
	(0.0009)	(0.0010)	(0.0009)	(0.0010)	(0.0009)	(0.0010)	(0.0010)	(0.0009)	(0.0010)	(0.0009)	(0.0010)	(0.0009)
TEXT_SPECIFICITY	−0.0001	−0.0002	−0.0002	−0.0002	−0.0001	−0.0001	0.0001	0.0001	0.0001	0.0000	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)

(*continued on next page*)

Table 5 (*continued*)

Subsample	High Institutional Ownership						Low Institutional Ownership					
Image Type	All		Specific		Generic		All		Specific		Generic	
Images Scaled By	Pages	Text	Pages	Text	Pages	Text	Pages	Text	Pages	Text	Pages	Text
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR	CAR
Control for Generic Image Usage	–	–	YES	YES	–	–	–	–	YES	YES	–	–
Control for Specific Image Usage	–	–	–	–	YES	YES	–	–	–	–	YES	YES
Firm Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	659	659	659	659	659	659	668	668	668	668	668	668
Adjusted R-squared	0.0908	0.0914	0.0922	0.0929	0.0896	0.0883	0.1135	0.1132	0.1114	0.1138	0.1132	0.1137

Table 6

ESG Report Image Usage and Shareholder Perception – Activist Proposals

This table reports OLS estimation of the model $PROPOSALS_{i,\,t+2} = \alpha + \beta_1 Image\ Usage_{i,t} + \beta_2 Image\ Usage_{i,t} * POOR\_ESG_{i,t} + \beta_3 Image\ Usage_{i,t} * SCORE\_CONTROVERSIES_{i,t} + \beta_4 POOR\_ESG_{i,t} + \beta_5 SCORE\_CONTROVERSIES_{i,t} + \beta_6 TEXT\_CONTENT_{i,t} + \beta_7 TEXT\_SPECIFICITY_{I,t} + Firm\ Controls_{i,t} + Industry\ Fixed\ Effects_i + Year\ Fixed\ Effects_t + \varepsilon$. PROPOSALS is the log-transformed number of shareholder proposals in the year after ESG report publication, i.e., t+2. Shareholders are required to submit proposals 120 days before the proxy statement date, which is usually on or after the annual report release date. Any proposals submitted for voting at the shareholder meeting in t+1 would've been submitted in t, before ESG reports for t are published in t+1. It is impossible for proposals at the t+1 shareholder meeting to be affected by ESG reports for t, and only possible for proposals recorded at the t+2 shareholder meeting to be affected by ESG reports for t.

Panel A reports results from regressions using all types of image usage as an independent variable. Panel B reports coefficients of interest (i.e., coefficients on specific image usage, the ESG performance measures, and their interactions) from regressions using specific image usage as an independent variable, while controlling for generic image usage. Panel C reports coefficients of interest from regressions using generic image usage as an independent variable, while controlling for specific image usage.

We split proposals into those sponsored by institutional investors in Columns 1 & 2 and those sponsored by individuals in Columns 3 & 4. Note that PROPOSALS is constructed on a firm basis and not on a proposal; hence, observations are 2111 firm-years throughout. For conciseness, *IMAGE_USAGE* alternately takes on the value of IMAGES/PAGE in columns labeled "Images Scaled By … Pages" or the value of IMAGES/TEXT in columns labeled "Images Scaled by … Text". Significance at the 0.01, 0.05 and 0.1 levels in two-tailed tests is denoted by ***, **, and *, respectively.

Panel A. Regressions on All Image Usage

Shareholder Group	Institutions (More Sophisticated)		Individuals (Less Sophisticated)	
Images Scaled by	Pages	Text	Pages	Text
VARIABLES	(1)	(2)	(3)	(4)
	PROPOSALS			
IMAGE_USAGE	0.2132**	0.0619**	0.0549	0.0288
	(0.0988)	(0.0249)	(0.1240)	(0.0316)
*IMAGE_USAGE**POOR_ESG	0.0002	−0.0001	0.0005	−0.0000
	(0.0021)	(0.0005)	(0.0027)	(0.0006)
*IMAGE_USAGE**SCORE_CONTROVERSIES	−0.0032***	−0.0009***	−0.0009	−0.0005*
	(0.0010)	(0.0003)	(0.0011)	(0.0003)
POOR_ESG	−0.0001	0.0001	−0.0010	−0.0007
	(0.0016)	(0.0014)	(0.0019)	(0.0017)
SCORE_CONTROVERSIES	0.0030***	0.0028***	0.0023***	0.0027***
	(0.0008)	(0.0007)	(0.0008)	(0.0007)
TEXT_CONTENT	−0.0001	−0.0001	0.0017	0.0015
	(0.0010)	(0.0010)	(0.0015)	(0.0015)
TEXT_SPECIFICITY	0.0000	0.0000	−0.0000	0.0000
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Firm Controls	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	2111	2111	2111	2111
Adjusted R-squared	0.3042	0.3047	0.3649	0.3660

Panel B. Coefficients of Interest from Regressions on Specific Image Usage

IMAGE_USAGE	0.2640*	0.0813**	0.1348	0.0564
	(0.1508)	(0.0396)	(0.1879)	(0.0497)
*IMAGE_USAGE**POOR_ESG	−0.0005	−0.0002	0.0012	0.0001
	(0.0030)	(0.0008)	(0.0037)	(0.0009)
*IMAGE_USAGE**SCORE_CONTROVERSIES	−0.0039**	−0.0011***	−0.0008	−0.0006
	(0.0015)	(0.0004)	(0.0016)	(0.0004)

(*continued on next page*)

Table 6 (*continued*)

POOR_ESG	0.0001	0.0001	−0.0011	−0.0008
	(0.0015)	(0.0014)	(0.0018)	(0.0017)
SCORE_CONTROVERSIES	0.0027***	0.0026***	0.0020***	0.0025***
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Panel C. Coefficients of Interest from Regressions on Generic Image Usage				
IMAGE_USAGE	0.5928	0.1689*	−0.2786	0.0123
	(0.3713)	(0.0868)	(0.4246)	(0.1006)
*IMAGE_USAGE**POOR_ESG	0.0025	0.0001	0.0039	0.0006
	(0.0072)	(0.0016)	(0.0086)	(0.0019)
*IMAGE_USAGE**SCORE_CONTROVERSIES	−0.0100***	−0.0027***	−0.0018	−0.0014*
	(0.0034)	(0.0008)	(0.0035)	(0.0008)
POOR_ESG	−0.0005	−0.0001	−0.0013	−0.0010
	(0.0016)	(0.0014)	(0.0018)	(0.0016)
SCORE_CONTROVERSIES	0.0028***	0.0027***	0.0020***	0.0025***
	(0.0007)	(0.0006)	(0.0007)	(0.0006)

Table 7

ESG Report Image Usage and ESG Awards

We estimate the following model using logistic regression: $ESG\_AWARD_{i,\ t+1} = \alpha + \beta_1 Image\ Usage_{i,t} + \beta_2 Image\ Usage_{i,t} * POOR\_ESG_{i,t} + \beta_3 Image\ Usage_{i,t} * SCORE\_CONTROVERSIES_{i,t} + \beta_4 POOR\_ESG_{i,t} + \beta_5 SCORE\_CONTROVERSIES_{i,t} + \beta_6 TEXT\_CONTENT_{i,t\ t} + \beta_7 TEXT\_SPECIFICITY_{i,t} + Firm\ Controls_{i,t} + Industry\ Fixed\ Effects_i + Year\ Fixed\ Effects_t + \varepsilon$. ESG_AWARD equals 1 if the company receives any ESG award in t+1 according to Refinitiv records and 0 otherwise. We exclude observations with unknown award status and observations with ESG reports released later than 1 year after fiscal year end.

For conciseness, *IMAGE_USAGE* alternately takes on the value of IMAGES/PAGE in columns labeled "Images Scaled By … Pages" or the value of IMAGES/TEXT in columns labeled "Images Scaled by … Text". Columns 1 & 2 use all object-detected image usage as an independent variable. Columns 3 & 4 focus on specific image usage as an independent variable while controlling for generic image usage. Columns 5 & 6 focus on generic image usage as an independent variable while controlling for specific image usage. Standard errors in parentheses are clustered at the firm level. Significance at the 0.01, 0.05 and 0.1 levels in two-tailed tests is denoted by ***, **, and *, respectively.

Image Type	All		Specific		Generic	
Images Scaled By	Pages	Text	Pages	Text	Pages	Text
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	ESG_AWARD	ESG_AWARD	ESG_AWARD	ESG_AWARD	ESG_AWARD	ESG_AWARD
IMAGE_USAGE	−1.2316	−0.3550	−1.6343	−0.3508	−5.0007*	−1.1948
	(0.8826)	(0.2250)	(1.3661)	(0.3654)	(2.9974)	(0.7307)
*IMAGE_USAGE**POOR_ESG	0.0185	0.0036	0.0360	0.0069	0.0297	0.0027
	(0.0166)	(0.0040)	(0.0241)	(0.0061)	(0.0519)	(0.0125)
*IMAGE_USAGE**SCORE_CONTROVERSIES	0.0151*	0.0038*	0.0251**	0.0059*	0.0391	0.0091
	(0.0084)	(0.0021)	(0.0123)	(0.0033)	(0.0271)	(0.0064)
POOR_ESG	−0.0524***	−0.0483***	−0.0548***	−0.0494***	−0.0454***	−0.0415***
	(0.0139)	(0.0135)	(0.0136)	(0.0133)	(0.0130)	(0.0126)
SCORE_CONTROVERSIES	−0.0028	−0.0017	−0.0036	−0.0018	−0.0002	0.0010
	(0.0062)	(0.0057)	(0.0062)	(0.0058)	(0.0057)	(0.0051)
TEXT_CONTENT	0.0375***	0.0340***	0.0377***	0.0347***	0.0377***	0.0348***
	(0.0103)	(0.0105)	(0.0103)	(0.0105)	(0.0104)	(0.0106)
TEXT_SPECIFICITY	0.0036*	0.0037*	0.0037*	0.0036*	0.0036*	0.0036*
	(0.0020)	(0.0019)	(0.0020)	(0.0019)	(0.0020)	(0.0020)
Control for Generic Image Usage	–	–	YES	YES	–	–
Control for Specific Image Usage	–	–	–	–	YES	YES
Firm Controls	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	1860	1860	1860	1860	1860	1860
Pseudo R2	0.182	0.180	0.185	0.183	0.183	0.182

Data availability

Data is publicly or commercially available as described in the paper. Code is available upon request.

References

Adukia, A., Eble, A., Harrison, E., Runesha, H. B., & Szasz, T. (2023). What we teach about race and gender: Representation in images and text of children's books. *Quarterly Journal of Economics, 138*(4), 2225–2285. https://doi.org/10.1093/qje/qjad028

Al-Tuwaijri, S. A., Christensen, T. E., & Hughes, K. E., II (2004). The relations among environmental disclosure, environmental performance, and economic performance: A simultaneous equations approach. *Accounting, Organizations and Society, 29*, 447–471.
Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal, 74*(3), 87–103.
Andrew, J., & Baker, M. (2020). Corporate social responsibility reporting: The last 40 years and a path to sharing future insights. *Abacus, 56*(1), 35–65.
Aouadi, A., & Marsat, S. (2018). Do ESG controversies matter for firm value? Evidence from international data. *Journal of Business Ethics, 151*, 1027–1047.
Arya, A., Glover, J., & Sunder, S. (1998). Earnings management and the revelation principle. *Review of Accounting Studies, 3*, 7–34.
Aswani, J., & Rajgopal, S. (2023). Rethinking the value and emission implications of green bonds. Working paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4215882.
Athey, S., Karlan, D., Palikot, E., & Yuan, Y. (2023). Smiles in profiles: Improving fairness and efficiency using estimates of user preferences in online marketplaces. Working paper https://www.nber.org/system/files/working_papers/w30633/w30633.pdf.
Avgerinou, M. D., & Petterson, R. (2011). Towards a cohesive theory of visual literacy. *Journal of Visual Literacy, 30*(2), 1–19.
Azevedo-Rezende, L., Makhoul, P., Musacchio, A., Pongeluppe, L. S., & Redkar, S. B. (2021). *Fifty shades of green: Greenwashing and accountability*. Working paper.
Bagnoli, M., & Watts, S. (2003). Selling to socially responsible consumers: Competition and the private provision of public goods. *Journal of Economics and Management Strategy, 12*, 419–445.
Baker, A., Larcker, D. F., McClure, C., Saraph, D., & Watts, E. M. (2022). Diversity washing. Working paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4298626.
Bansal, P., & Clelland, I. (2004). Talking trash: Legitimacy, impression management, and unsystematic risk in the context of the natural environment. *Academy of Management Journal, 47*(1), 93–103.
Bansal, P., & Kistruck, G. (2006). Seeing is (not) believing: Managing the impression of the firm's commitment to the natural environment. *Journal of Business Ethics, 67*, 165–180.
Barton, J., & Mercer, M. (2005). To blame or not to blame: Analysts' reactions to explanations of poor management performance. *Journal of Accounting and Economics, 39*, 509–533.
Bartov, E., Radhakrishnan, S., & Krinsky, I. (2000). Investor sophistication and patterns in stock returns after earnings announcements. *The Accounting Review, 75*, 43–63.
Bartov, E., Givoly, D., & Hayn, C. (2002). The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics, 33*, 173–204.
Beattie, V., & Jones, M. J. (1992). The use and abuse of graphs in annual reports: Theoretical framework and empirical study. *Accounting and Business Research, 22*, 291–303.
Beattie, V., & Jones, M. J. (2000a). Changing graph use in corporate annual reports: A time-series analysis. *Contemporary Accounting Research, 17*(2), 213–226.
Beattie, V., & Jones, M. J. (2000b). Impression management: The case of inter-country financial graphs. *Journal of International Accounting, Auditing and Taxation, 9*(2), 159–183.
Beattie, V., & Jones, M. (2008). Corporate reporting using graphs: A review and synthesis. *Journal of Accounting Literature, 27*, 71–110.
Ben-Raphael, A., Ronen, J., Ronen, T., & Zhou, M. (2023). "Show me!" the informativeness of images. Working paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3954219.
Benschop, Y., & Meihuizen, H. E. (2002). Keeping up gendered appearances: Representations of gender in financial annual reports. *Accounting, Organizations and Society, 27*(7), 611–636.
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
Boiral, O. (2013). Sustainability reports as simulacra? A counter-account of A and A+ GRI reports. *Accounting, Auditing & Accountability Journal, 26*(7), 1036–1071.
Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management, 34*(6), 1080–1109.
Bonetti, P., Cho, C. H., & Michelon, G. (2023). Environmental disclosure and the cost of capital: Evidence from the Fukushima nuclear disaster. *European Accounting Review*. https://doi.org/10.1080/09638180.2023.2203410
Bothello, J., Ioannou, I., Porumb, V., & Zengin-Karaibrahimoglu, Y. (2023). CSR decoupling within business groups and the risk of perceived greenwashing. *Strategic Management Journal, 44*(13), 3217–3251.
Bowen, R. M., Davis, A. K., & Matsumoto, D. A. (2005). Emphasis on pro forma versus GAAP earnings in quarterly press releases: Determinants, SEC intervention and market reactions. *The Accounting Review, 80*(4), 1011–1038.
Bushee, B. (1998). The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review, 73*, 305–333.
Cai, Y., Jo, H., & Pan, C. (2012). Doing well while doing bad? ESG in controversial industry sectors. *Journal of Business Ethics, 108*(4), 467–480.
Cao, S., Cheng, Y., Wang, M., Xia, Y., & Yang, B. (2023). Visual information in the age of AI: Evidence from corporate executive presentations. Working paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4490834.
Cerrato, D., & Ferrando, T. (2020). The financialization of civil society activism: Sustainable finance, non-financial disclosure and the shrinking space for engagement. *Accounting, Economics, and Law: Convivium*. https://doi.org/10.1515/ael-2019-0006
Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? *Journal of Economics and Management Strategy, 18*(1), 125–169.
Cho, C. H., Michelon, G., & Patten, D. M. (2012). Impression management in sustainability reports: An empirical investigation of the use of graphs. *Accounting and the Public Interest, 12*, 16–37.
Cho, C. H., Phillips, J. R., Hageman, A. M., & Patten, D. M. (2009). Media richness, user trust, and perceptions of corporate social responsibility: An experimental investigation of visual web site disclosures. *Accounting, Auditing & Accountability Journal, 22*(6), 933–952.
Christensen, H. B., Floyd, E., Liu, L. Y., & Maffett, M. (2017). The real effects of mandated information on social responsibility in financial reports: Evidence from mine-safety records. *Journal of Accounting and Economics, 64*, 284–304.
Christensen, T. E., Fronk, K. E., Lee, J. A., & Nelson, K. K. (2024). Data visualization in 10-K filings. *Journal of Accounting and Economics, 77*(2–3), 1–34.
Christensen, H., Hail, L., & Leuz, C. (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Review of Accounting Studies, 26*, 1176–1248.
Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review, 97*(1), 147–175.
Cohn Reznick. (2023). Meet the 2023 judges. https://www.cohnreznick.com/insights/meet-the-judges-esg-gamechangers. (Accessed 10 October 2024).
Cooper, E. W., & Uzun, H. (2015). Corporate social responsibility and the cost of debt. *Journal of Accounting and Finance, 15*(8), 11–29.
Courtis, J. K. (2004). Colour as visual rhetoric in financial reporting. *Accounting Forum, 28*, 265–281.
Darendeli, A., Fiechter, P., Hitz, J., & Lehmann, N. (2022). The role of corporate social responsibility (CSR) information in supply-chain contracting: Evidence from the expansion of CSR rating coverage. *Journal of Accounting and Economics, 74*, 1–26.
De Villiers, C., Jia, J., & Li, Z. (2022). Corporate social responsibility: A review of empirical research using thomson reuters Asset4 data. *Accounting and Finance, 62*(4), 4523–4568.
Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics, 50*(2–3), 344–401.
Deng, W., Gao, L., Hu, B., & Zhou, G. (2023). Seeing is believing: Annual report "graphicity" and stock returns. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3723126.
Desai, H., Lam, P., Lin, B., & Rajgopal, S. (2023). An analysis of carbon-reduction pledges of U.S. oil and gas companies. *Management Science, 69*(6), 3748–3758.
Dewatripoint, M., & Tirole, J. (2005). Modes of communication. *Journal of Political Economy, 113*(6), 1217–1238.
Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate responsibility reporting. *The Accounting Review, 86*(1), 59–100.
Diouf, D., & Boiral, O. (2017). The quality of sustainability reports and impression management: A stakeholder perspective. *Accounting, Auditing & Accountability Journal, 30*(3), 643–667.
Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review, 20*(1), 65–91.
Dorfleitner, G., Kreuzer, C., & Sparrer, C. (2020). ESG controversies and controversial ESG: About silent saints and small sinners. *Journal of Asset Management, 21*, 393–412.
Dowling, J., & Pfeffer, J. (1975). Organizational legitimacy: Social values and organizational behavior. *Pacific Sociological Review, 18*(1), 122–136.
Dupire, M., & M'Zali, B. (2018). ESG strategies in response to competitive pressures. *Journal of Business Ethics, 148*, 603–623.
Elliott, W. B. (2006). Are investors influenced by pro forma emphasis and reconciliations in earnings announcements? *The Accounting Review, 81*(1), 113–133.
Elliott, W. B., Grant, S. M., & Rennekamp, K. M. (2017). How disclosure features of corporate social responsibility reports interact with investor numeracy to influence investor judgements. *Contemporary Accounting Research, 34*(3), 1596–1621.
Elliott, W. B., Jackson, K. E., Peecher, M. E., & White, B. J. (2014). The unintended effect of corporate social responsibility performance on investors' estimates of fundamental value. *The Accounting Review, 89*(1), 275–302.
Elsbach, K. D., & Sutton, R. I. (1992). Acquiring organizational legitimacy through illegitimate actions: A marriage of institutional and impression management theories. *Academy of Management Journal, 35*(4), 699–738.
ESG & Sustainability Awards. (2024). Judges. https://www.esg-awards.com/judges. (Accessed 10 October 2024).
ESG Investing: Sustainability News, Events & Awards. (2024). Carbon awards 2024 judging panel. https://www.esginvesting.co.uk/awards-carbon-judges/. (Accessed 10 October 2024). Published September 20, 2024.
European Commission. (2013). *Proposal for a directive of the European parliament and of the council amending council directives 78/660/EEC and 83/349/EEC as regards disclosure of non-financial and diversity information by certain large companies and groups*. COM(2013) 207 final, 2.
European Commission. (2021). Sustainable finance package. Published April 21, 2021 https://ec.europa.eu/info/publications/210421-sustainable-finance-communication_en. (Accessed 14 June 2021).
Fabozzi, F. J., Ma, K. C., & Oliphant, B. J. (2008). Sin stock returns. *Journal of Portfolio Management, 34*(4), 82–94.
Fama, E. F., & French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics, 43*, 153–159.
Farrell, J., & Rabin, M. (1996). Cheap talk. *The Journal of Economic Perspectives, 10*(3), 103–118.

Finkel, J. R., Grenager, T., & Manning, C. (2005). Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd annual meeting on association for computational linguistics*.

Flammer, C. (2015a). Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. *Management Science, 61*(11), 2549–2568.

Flammer, C. (2015b). Does product market competition foster corporate social responsibility? Evidence from trade liberalization. *Strategic Management Journal, 36*, 1469–1485.

Flammer, C., Toffel, M. W., & Viswanathan, K. (2021). Shareholder activism and firms' voluntary disclosure of climate change risks. *Strategic Management Journal, 42*(10), 1850–1879. https://doi.org/10.1002/smj.3313

Frankental, P. (2001). Corporate social responsibility – A PR invention? *Corporate Communications, 6*(1), 18–23.

Frederickson, J. R., & Miller, G. S. (2004). The effects of pro forma earnings disclosure on analysts' and nonprofessional investors' equity valuation judgments. *The Accounting Review, 79*(3), 667–686.

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston: Pitman.

Gamerschlag, R., Möller, K., & Verbeeten, F. (2011). Determinants of voluntary ESG disclosure: Empirical evidence from Germany. *Review of Managerial Science, 5*(2–3), 233–262.

Gantchev, N., & Gianetti, M. (2020). Costs and benefits of shareholder democracy: Gadflies and low-cost activism. *Review of Financial Studies, 34*(12), 5629–5675.

García-Sánchez, I. M., & Araújo-Bernardo, C. A. (2019). What colour is the corporate social responsibility report? Structural visual rhetoric, impression management strategies, and stakeholder engagement. *Corporate Social Responsibility and Environmental Management, 27*(2), 1117–1142.

Global Reporting Initiative. (2020). About GRI. https://www.globalreporting.org/. (Accessed 28 March 2020).

Godfrey, J., Mather, P., & Ramsay, A. (2003). Earnings and impression management in financial reports: The case of CEO changes. *Abacus, 39*(1), 95–123.

Godfrey, P. C., Merrill, C. B., & Hansen, J. M. (2009). The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. *Strategic Management Journal, 30*, 425–445.

Good Jobs First. (2015). Introducing violation tracker. https://goodjobsfirst.org/introducing-violation-tracker. (Accessed 20 August 2023).

Goodman, J. (2015). Religious organizations as shareholders: Salience and empowerment. In M. Goranova, & L. V. Ryan (Eds.), *Shareholder empowerment: A new era in corporate governance*. New York: Palgrave Macmilan.

Graphic Artists Guild. (2021). *Graphic artists Guild handbook: Pricing & ethical guidelines* (16th ed.). Cambridge, Massachusetts: MIT Press.

Graves, O. F., Flesher, D. L., & Jordan, R. E. (1996). Pictures and the bottom line: The television epistemology of U.S. annual reports. *Accounting, Organizations and Society, 21*(1), 57–88.

Grewal, J., Riedl, E. J., & Serafeim, G. (2019). Market reaction to mandatory nonfinancial disclosure. *Management Science, 65*(7), 3061–3084.

Grougiou, V., Dedoulis, E., & Leventis, S. (2016). Corporate social responsibility reporting and organizational stigma: The case of "sin" industries. *Journal of Business Research, 69*(2), 905–914.

Hancock, P. (2003). *Aestheticizing the world of organization—creating beautiful untrue things. Art and Aesthetics at work*. New York: Palgrave MacMillan.

Hand, J. (1990). A Test of the extended functional fixation Hypothesis. *The Accounting Review, 65*, 740–763.

Heese, J., Perez-Cavazos, G., & Peter, C. D. (2022). When the local newspaper leaves town: The effects of local newspaper closures on corporate misconduct. *Journal of Financial Economics, 145*(2), 445–463.

Heflin, F., & Wallace, D. (2017). The BP oil spill: Shareholder wealth effects and environmental disclosures. *Journal of Business Finance & Accounting, 44*(3–4), 337–374.

Herring, J. (1990). *Annual report design: A guide to the annual report process for graphic designers and corporate communicators*. New York: Watson-Guptill.

Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy, 124*(5), 1423–1465. http://hobergphillips.tuck.dartmouth.edu/industryconcern.htm.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics, 93*(1), 15–36.

Hooghiemstra, R. (2000). Corporate communication and impression management—New perspectives why companies engage in corporate social reporting. *Journal of Business Ethics, 27*, 55–68.

Hope, O.-K., Hu, D., & Lu, H. (2016). The benefits of specific risk-factor disclosures. *Review of Accounting Studies, 21*(4), 1005–1045.

Hopwood, A. G. (2009). Accounting and the environment. *Accounting, Organizations and Society, 34*, 433–439.

Hrasky, S. (2012). Visual disclosure strategies adopted by more or less sustainability-driven companies. *Accounting Forum, 36*, 154–165.

Huang, X., Teoh, S. H., & Zhang, Y. (2014). Tone management. *The Accounting Review, 89*(3), 1083–1113.

IJGlobal. (2023). IJGlobal ESG awards 2023 – The Judging panel (so far). https://www.ijglobal.com/articles/173726/ijglobal-esg-awards-2023-the-judging-panel-so-far. (Accessed 10 October 2024). Published July 11, 2023.

Ioannou, I., Kassinis, G., & Papagiannakis, G. (2022). The impact of perceived greenwashing on customer satisfaction and the contingent role of capability reputation. *Journal of Business Ethics, 185*(2), 333–347.

Ioannou, I., & Serafeim, G. (2015). The impact of corporate social responsibility on investment recommendations: Analysts' perceptions and shifting institutional logics. *Strategic Management Journal, 36*(7), 1053–1081.

Jensen, M. C. (2002). Value maximization, stakeholder theory, and the corporate objective function. *Business Ethics Quarterly, 12*(2), 235–256.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305–360.

Jones, M. J. (2011). The nature, use, and impression management of graphs in social and environmental accounting. *Accounting Forum, 35*, 75–89.

Kirk, M. P., & Vincent, J. D. (2014). Professional investor relations within the firm. *The Accounting Review, 89*, 1421–1452.

Kitzmueller, M., & Shimshack, J. (2012). Economic perspectives on corporate social responsibility. *Journal of Economic Literature, 50*(1), 51–84.

Knobloch, S., Hastall, M., Zillmann, D., & Callison, C. (2003). Imagery effects on the selective reading of internet newsmagazines. *Communication Research, 30*(1), 3–29.

Koh, P., Qian, C., & Wang, H. (2014). Firm litigation risk and the insurance value of corporate social performance. *Strategic Management Journal, 35*, 1464–1482.

Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics, 60*(2–3), 110–135.

Larcker, D. F., & Watts, E. M. (2020). Where's the greenium? *Journal of Accounting and Economics, 69*(2), Article 101312.

Leonelli, S., Muhn, M., Rauter, T., & Sran, G. (2023). How do consumers use firm disclosure? Evidence from a randomized field experiment. In *Working paper. Proceedings, 2023 NYU big apple Accounting conference*.

Lev, B., Petrovits, C., & Radhakrishnan, S. (2010). Is doing good good for you? How corporate charitable contributions enhance revenue growth. *Strategic Management Journal, 31*, 182–200.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics, 45*, 221–247.

Li, Y., Gong, M., Zhang, X., & Koh, L. (2018). The impact of environmental, social, and governance disclosure on firm value: The role of CEO power. *The British Accounting Review, 50*, 60–75.

Li, Q., Watts, E. M., & Zhu, C. (2023). Retail investors and ESG news. Working paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4384675Lins.

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance, 72*(4), 1785–1824.

Lo, K. (2008). Earnings management and earnings quality. *Journal of Accounting and Economics, 45*, 350–357.

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research, 40*, 1–44.

Ludwig, J., & Mullainathan, S. (2024). Machine learning as a tool for hypothesis generation. *Quarterly Journal of Economics, 139*(2), 751–827. https://doi.org/10.1093/qje/qjad055

Lys, T., Naughton, J. P., & Wang, C. (2015). Signaling through corporate accountability reporting. *Journal of Accounting and Economics, 60*(1), 56–72.

Manchiraju, H., & Rajgopal, S. (2017). Does corporate social responsibility (CSR) create shareholder value? Evidence from the Indian companies act 2013. *Journal of Accounting Research, 55*(5), 1257–1300.

Martin, P. R., & Moser, D. V. (2016). Managers' green investment disclosures and investors' reaction. *Journal of Accounting and Economics, 61*(1), 239–254.

Mather, D., Mather, P., & Ramsay, A. (2005). An investigation into the measurement of graph distortion in financial reports. *Accounting and Business Research, 35*(2), 147–159.

McWilliams, A., & Siegel, D. S. (2000). Corporate social responsibility and financial performance: Correlation or misspecification? *Strategic Management Journal, 21*, 603–609.

McWilliams, A., & Siegel, D. S. (2001). Corporate social responsibility: A theory of the firm perspective. *Academy of Management Review, 26*(1), 117–127.

McWilliams, A., Siegel, D. S., & Wright, P. M. (2006). Guest editors' introduction. Corporate social responsibility: Strategic implications. *Journal of Management Studies, 43*(1), 1–18.

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.

Milne, M. J., & Grubnic, S. (2011). Climate change accounting research: Keeping it interesting and different. *Accounting, Auditing & Accountability Journal, 24*(8), 948–977.

Minor, D., & Morgan, J. (2011). ESG as reputation insurance: Primum non nocere. *California Management Reivew, 53*(3), 40–59.

Moneva, J. M., Archel, P., & Correa, C. (2006). GRI and the camouflaging of corporate unsustainability. *Accounting Forum, 30*(2), 121–137.

Moss, A., Naughton, J. P., & Wang, C. (2024). The irrelevance of environmental, social, and governance disclosure to retail investors. *Management Science, 70*(4), 2626–2644.

Muino, F., & Trombetta, M. (2009). Does graph disclosure bias reduce the cost of equity capital? *Accounting and Business Research, 39*(2), 83–102.

Nekrasov, A., Teoh, S. H., & Wu, S. (2021). Visuals and attention to earnings news on Twitter. *Review of Accounting Studies*. https://doi.org/10.1007/s11142-021-09630-8

Neu, D., Warsame, H., & Pedwell, K. (1998). Managing public impressions: Environmental disclosures in annual reports. *Accounting, Organizations and Society, 23*(3), 265–282.

Obaid, K., & Pukthuanthong, K. (2022). A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. *Journal of Financial Economics, 144*, 273–297.

Pacheco-Ortiz, D., Escobar-Sierra, M., & Suárez-Monsalve, A.-M. (2024). Corporate hypocrite in the context of corporate social responsibility: A literature review. *Social Responsibility Journal, 20*(8), 1576–1596.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*(3), 241.
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255–287.
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.
Penrose, J. M. (2008). Annual report graphic use: A review of the literature. *Journal of Business Communication, 45*(2), 158–180.
Preston, A. M., Wright, C., & Young, J. J. (1996). IMag[in]ing annual reports. *Accounting, Organizations and Society, 21*(1), 113–137.
Raghunandan, A., & Rajgopal, S. (2023). Do socially responsible firms walk the talk? *Journal of Law and Economics, 67*(4), 767–810.
Refinitiv. (2019). Environmental, social and governance (ESG) scores from refinitiv. https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf. (Accessed 29 March 2020).
Reitmaier, C., Schultze, W., & Vollmer, J. (2024). Corporate responsibility and corporate misbehavior: Are CSR reporting firms indeed responsible? *Review of Accounting Studies*. https://doi.org/10.1007/s11142-024-09850-8. online preprint.
Reverte, C. (2012). The impact of better corporate social responsibility disclosure on the cost of equity capital. *Corporate Social Responsibility and Environmental Management, 19*, 253–272.
Ronen, J., Ronen, T., Zhou, M. J., & Gans, S. E. (2023). The informational role of imagery in financial decision making: A new approach. *Journal of Behavioral and Experimental Finance, 40*, 1–30.
SASB Symposium Roundtable. (2016). The next wave of ESG integration: Lessons from institutional investors. *Journal of Applied Corporate Finance, 29*(2), 32–44.
Schadewitz, H., & Niskala, M. (2010). Communication via responsibility reporting and its effect on firm value in Finland. *Corporate Social Responsibility and Environmental Management, 17*, 96–106.
SEC. (2022). SEC charges BNY mellon investment adviser for misstatements and omissions concerning ESG considerations. https://www.sec.gov/news/press-release/2022-86. (Accessed 4 September 2024).
Serafeim, G., & Yoon, A. (2022). Which corporate ESG news does the market react to? *Financial Analysts Journal, 78*(I), 59–78.
Shabana, K. M., & Ravlin, E. C. (2016). Corporate social responsibility as substantive and symbolic behavior: A multilevel theoretical analysis. *Business and Society Review, 121*(2), 297–327.
Sobel, J. (2015). Signaling games. In R. A. Meyers (Ed.), *Encyclopedia of complexity and systems science*. New York, New York: Springer Science+Business Media. https://doi.org/10.1007/978-3-642-27737-5_481-4.
Soltes, E. (2019). The frequency of corporate misconduct: Public enforcement versus private reality. *Journal of Financial, Crime*, 923–937.
Spiller, R. (2000). Ethical business and investment: A model for business and society. *Journal of Business Ethics, 27*(1/2), 149–160. https://doi.org/10.1023/A:1006445915026
Steinbart, P. J. (1989). The auditor's responsibility for the accuracy of graphs in annual reports: Some evidence of the need for additional guidance. *Accounting Horizons, 3*, 60–70.
Taylor, B. G., & Anderson, L. K. (1986). Misleading graphs-guidelines for the accountant. *Journal of Accountancy, 162*(4), 126.
Thomson Reuters. (2018). Thomson reuters financial & risk business announces new company name: Refinitiv. https://www.thomsonreuters.com/en/press-releases/2018/july/thomson-reuters-financial-and-risk-business-announces-new-company-name-refinitiv.html. (Accessed 29 March 2020).
Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics, 5*, 179–194.
Voth, H., & Yanagizawa-Drott, D. (2024). Image(s). Working paper https://www.jvoth.com/images_july_24.pdf.
Walther, B. (1997). Investor sophistication and market earnings expectations. *Journal of Accounting Research, 35*, 157–192.
WRDS. (2020). https://wrds-www.wharton.upenn.edu/pages/wrds-research/applications/institutional-ownership-research/institutional-ownership-concentration-and-breadth-ratios/.
Yermack, D. (2010). Shareholder voting and corporate governance. *Annual Review of Financial Economics, 2*, 103–125.
Zillmann, D., Gibson, R., & Sargent, S. L. (1999). Effects of photographs in news-magazine reports on issue perception. *Media Psychology, 1*, 207–228.