Graham Bullock*

Signaling the credibility of private actors as public agents: transparency, independence, and expertise in environmental evaluations of products and companies

Abstract: Private firms are increasingly being regarded as moral agents of their stakeholders and the broader public. Stakeholders can use different types of evaluation organizations to monitor this division of moral labor, but must also monitor the credibility of this second layer of moral agents. This paper uses agency, signaling, and legitimacy theory to develop a novel conceptual framework showing how both firms and evaluation organizations send signals of their credibility as moral agents to earn grants of legitimacy from their moral stakeholders. The paper also describes how three specific characteristics of ratings and certifications – transparency, expertise, and independence – may signal different forms of credibility, appeal to particular stakeholder groups, and elicit different forms of legitimacy. A content analysis of the websites of 245 eco-labels, sustainability ratings and other forms of environmental evaluations reveals the multi-dimensional nature of these three characteristics, and finds that transparency is the most commonly-sent signal of credibility, followed by independence and then expertise. These results highlight the complexity of existing signals of credibility, and suggest several strategies – including voluntary credibility standards and a virtual information marketplace – that both private and public actors can pursue to improve the quality and accessibility of these signals of credibility.

DOI 10.1515/bap-2014-0028
Previously published online March 28, 2015

***Corresponding author: Graham Bullock,** Political Science and Environmental Studies, Davidson College, Box 7134, Davidson, NC 28035-7134, USA, e-mail: grbullock@davidson.edu.
http://orcid.org/0000-0002-3680-206X

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

1 Introduction

Milton Friedman famously stated that the "social responsibility of business is to increase its profits."[1] Alternatively, stakeholder, stewardship and principle theory assert that executives are not only ethically obligated to take into account the interests of shareholders but also those of employees, community members and other stakeholders.[2] Such conceptions of corporate social responsibility and extended corporate citizenship have increasingly created a "division of moral labor" that privatizes "the major public responsibilities of liberal democratic society."[3] This division of labor suggests that corporations should also serve as "moral agents" of their stakeholders, and that these stakeholders need effective mechanisms to monitor this broader principal-agent relationship.[4]

Such oversight can be based on information produced by the firms themselves or on externally-produced evaluations of companies and their products, which can include certifications, databases, ratings, reviews, awards, and boycotts.[5] For example, consumers can use environmental certifications such as USDA Organic to identify products that are safer for their health and the environment,[6] employees can use rankings such as Fortune's Best Places to Work to compare employment opportunities,[7] and investors can use indices such as the Dow Jones Sustainability Index to invest in companies that are socially responsible.[8] Likewise, firms such as Hershey's can rely on non-profit organizations such as Rainforest Alliance to certify that its suppliers are sourcing their chocolate from environmentally and socially responsible farms.[9]

While such delegation of monitoring can have important efficiency advantages, it also creates a second principal-agent problem for stakeholders.[10] As in the case of firms acting as agents of public responsibility, the incentives of the organizations behind these externally-produced environmental evaluations

1 Milton Friedman (1970: p. SM17).
2 Caldwell, Karri, and Vollmar (2006).
3 Mäkinen and Kourula (2012: p. 666).
4 Ludescher, Mahsud, and Prussia (2012).
5 Boiral and Gendron (2011). See Appendix 1 for definitions of these different types of evaluations. For clarity and concision, the paper will henceforth refer to these phenomena as "environmental evaluations," "programs," or "initiatives."
6 Kiesel and Villas-Boas (2007).
7 Turban and Cable (2003).
8 López, Garcia, and Rodriguez (2007).
9 See <http://www.hersheys.com/bliss/our-story/rainforest-alliance.aspx>. Accessed November 17, 2014.
10 Starobin and Weinthal (2010); Lin (2012).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

(henceforth, referred to as "evaluation organizations") may be imperfectly aligned with the interests of their stakeholder principals, regardless of whether they are non-profit, governmental, or for-profit institutions. Traditional agency theory suggests that principals can control their agents by providing them with either strict rules to follow or broad principles to interpret. Principals can monitor adherence to these rules and principles by either proactively implementing "police patrols" or efficiently responding to "fire alarms" set off by third parties.[11]

Consumer, employee, and community stakeholders, however, often lack the power to utilize these control mechanisms. They therefore often depend on more indirect forms of regulation, such as reputational pressures and the granting of legitimacy,[12] which may not be as effective, especially when their exit options are limited. This principal-agent problem is exacerbated by the fact that sustainability information is usually a "credence good." Such information requires trust in its quality even after its use because it is difficult to know how accurate it is.[13] This partially explains why 56% of Americans do not trust companies' green claims.[14] Starobin and Weinthal's analysis of the Kosher label demonstrates how transparency, expertise and verification grounded in social networks can be particularly important to the establishment of such trust and credibility, and challenges the value of claims of "third party independence."[15]

This paper builds on Starobin and Weinthal's work by examining how evaluation efforts beyond the Kosher label send signals of credibility to the public. To better understand these signals, the first part of the paper presents a novel conceptual framework that connects agency and signaling theory with the concepts of accountability, credibility, and legitimacy. The second part discusses three particular characteristics – transparency, independence, and expertise – that often serve as signals to principals, and suggests an important empirical question – which of these characteristics are most commonly used by firms and evaluation organizations to communicate their credibility? In order to address this question, the third part of the paper presents data from original research on 245 environmental evaluations of products and companies, and finds that transparency is the most commonly-found attribute, followed by independence and then expertise. The paper's final section discusses the managerial, societal and theoretical implications of this research, and outlines the opportunities for policy innovation

11 McCubbins and Schwartz (1984); Lin (2012).

12 Lin (2012); Deegan (2002).

13 Nelson (1970); Darby and Karni (1973); Nadaï (1999); Roe and Sheldon (2007).

14 See <http://www.conecomm.com/2012-cone-green-gap-trend-tracker>. Accessed August 25, 2013.

15 Starobin and Weinthal (2010).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

and future research that it presents. The paper concludes with a summary of the relevance of these results to the field of business and politics and its understanding of private regulation.

2 Conceptual framework: signals and accountability relationships

This section introduces a novel conceptual framework for understanding the relationship between providers of sustainability information ("moral agents") and receivers of that information ("moral principals"). The framework uses signaling and agency theory to connect the concepts of accountability, credibility, and legitimacy, terms that are frequently used to describe eco-labels in the literature but are often not clearly defined nor are the relationships between them clearly explained. The concept of accountability is particularly useful in bringing additional clarity to the principal-agent relationships described above. While accountability is conventionally thought of as "responsibility for a set of defined, concrete assets," it is better conceived of as "being accountable to a set of internal and external stakeholders."[16] As Black explains, accountability is a type of relationship "between different actors in which one gives account and another has the power or authority to impose consequences."[17]

Such giving of account relates to the economic theory of signaling, which posits that actors can solve the problem of information asymmetry between them by sharing relevant information about themselves.[18] Agency theory further posits that such information can be used by principals to control agent opportunism.[19] In the context of sustainability, this process involves firms – as "moral agents" of their stakeholders – demonstrating the extent to which they meet their "social, environmental and economic responsibilities through compliance with standards and established rules in the area."[20] As Figure 1 shows, firms can signal such accountability directly themselves through advertisements, sustainability reports and their own eco-labels or indirectly via certifications and other forms of assessment produced by external evaluation organizations.[21] Thus either these

16 Painter-Morland (2007: p. 515–526).
17 Black (2008: p. 150).
18 Spence (1973).
19 Eisenhardt (1989).
20 Boiral and Gendron (2011: p. 332).
21 Boiral and Gendron (2011).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

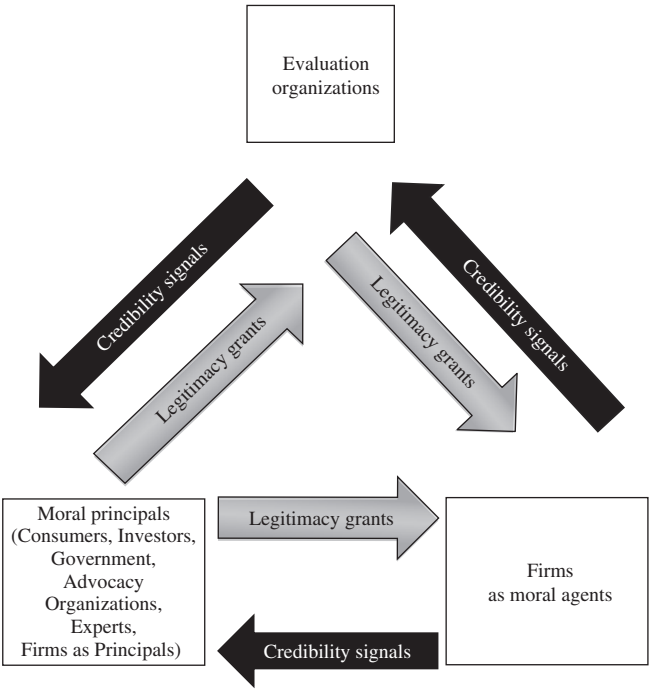

Figure 1: A three-way accountability relationship.
Credibility signals can include attributes demonstrating qualification and expertise, safety and trustworthiness, dynamism, and outcome effectiveness, while *legitimacy grants* can include the bestowal or withdrawal of economic resources, social or political status, information, authority and influence, and physical force.

evaluation organizations or firms publicizing their performance directly can serve as signalers and "moral agents."[22] In order to serve as useful information for their "moral principals" who are concerned about sustainability, signals from these agents must be less costly for firms that have more sustainable products and operations.

Upon receiving and evaluating these signals, principals can impose either positive or negative consequences on the agents sending these signals. These consequences are created when stakeholders deploy the different resources of power they have at their command, which include economic resources, social

22 As Figure 1 shows, firms can also serve as moral principals when they are clients or customers in business-to-business relationships.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

and political status, information, authority and influence, and physical force.[23] For example, socially responsible investors often choose to invest more in companies that they assess as sending credible signals,[24] while activists have attacked the social or political status of firms and evaluation organizations they do not find credible.[25] Consumers are generally less willing to share information about their preferences with firms that do not send signals of trustworthiness,[26] while government regulators are often more flexible with firms with trustworthy reputations.[27] Firms acting as principals in business-to-business relationships are more likely to trust suppliers that report strong corporate social responsibility. In extreme cases, low credibility firms that violate salient norms have suffered from applications of physical force, such as boycotts, sit-ins, and attacks on facilities.[28]

These consequences can be understood as themselves signaling a principal's assessment of an agent's legitimacy. As Figure 1 illustrates, accountability is therefore a three-way relationship in which moral principals bestow or withdraw grants of legitimacy based on the signals of credibility sent by their moral agents (either evaluation organizations or firms), while evaluation organizations may also bestow or withdraw grants of legitimacy to firms based on their assessment processes.[29] The sections below describe these signals of credibility and grants of legitimacy in more detail.

2.1 Credibility: what to signal?

The concept of credibility can help identify the types of characteristics that might serve as effective signals to principals. Credibility has been defined as "the

23 Ilchman and Uphoff (1969).
24 Mercer (2004).
25 Rowley and Berman (2000).
26 McKnight, Choudhury, and Kacmar (2002).
27 Potoski and Prakash (2004).
28 Monroe Friedman (1999); Liddick (2006).
29 It is possible for each of these legitimacy grants and credibility signals to flow in the reverse direction. For example, evaluation organizations may send credibility signals to firms in order to attract their business, while firms may bestow legitimacy grants to advocacy organizations by partnering with them. While beyond the scope of this paper, such reverse flows may be significant in certain contexts and represent an important area of future research.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

quality or power of inspiring belief,"[30] or as believability or authoritativeness.[31] This definition is instructive because it suggests that accepting a claim's credibility is to "take it on faith," even absent more tangible and direct evidence of outcome effectiveness.

Social science research has identified several different dimensions of credibility. One of the most influential frameworks was developed by Hovland, Janis, and Kelley and identifies two primary dimensions of credibility – perceived expertise and trustworthiness.[32] Berlo, Lemert, and Mertz later identified three related dimensions – safety, qualification, and dynamism.[33] Safety includes not only trustworthiness but also the perceived intent of the source and its altruism, fairness, calmness, patience, friendliness, and kindness. Similarly, qualification connects to Hovland, Janis and Kelley's perceived expertise, and includes factors such as training, experience, and knowledge about the relevant subject as well as more general ability, intelligence, and authoritativeness.[34] Dynamism, which includes boldness, frankness, and energy, is a potentially independent variable that can act as an intensifier of the other factors.

As many scholars have pointed out, credibility is a relational concept, and must be understood in relation to the perceptions of relevant stakeholders.[35] Thus the characteristics that agents choose as signals may be more credible to some stakeholder groups than others. This raises the important possibility that agents may be sending particular signals in order to attract the support of specific stakeholders. Given this potential dynamic, it is important to consider the underlying motivations and the role of legitimacy in driving these signaling processes.

2.2 Legitimacy: signaling for what purpose?

Past research has identified several reasons why firms decide to disclose information about their environmental and social performance. Such disclosure may be motivated by perceived business advantages or ethical responsibilities, existing or potential legislation, borrowing requirements, community and stakeholder expectations, employee needs, and legitimacy threats.[36] Legitimacy concerns may

30 "Credibility" (2014).
31 Fogg and Tseng (1999); Cash et al. (2002).
32 Hovland, Janis, and Kelley (1953).
33 Berlo, Lemert, and Mertz (1969).
34 Hovland, Janis, and Kelley (1953).
35 Shapin (1995); Boström (2006); Alagona (2008).
36 Adams, Hill, and Roberts (1998); Deegan (2002); Farneti and Guthrie (2009); Paulraj (2009).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

be particularly strong motivations; for example, environmental disclosures in oil company annual reports more than doubled after their legitimacy was threatened by the Exxon Valdez oil spill.[37] Vogel suggests that while some firms signal their social responsibility in order to minimize risks to their reputation, others do so because it is part of their "corporate strategy and business identity."[38]

Suchman defines legitimacy as the belief that "the actions of an entity are desirable, proper, or appropriate,"[39] while Uphoff describes how legitimacy is granted to individuals or organizations "in keeping with the beliefs people have about what is right and proper."[40] Legitimacy theory suggests that organizations depend on legitimacy for their survival and will use strategies such as information disclosure to ensure its continued supply, while stakeholder theory further suggests that organizations will disclose information that is salient to stakeholders they perceive as particularly important sources of legitimacy.[41]

Such disclosures can earn an organization several different types of legitimacy. Information that contributes to the self-interest of stakeholders can enhance an organization's *pragmatic legitimacy*, information that enhances the welfare of society can enhance its *moral legitimacy*, and information that encourages stakeholders to view an organization as a natural and inevitable part of their lives can enhance its *cognitive legitimacy*.[42] Cashore demonstrates the relevance of this framework to analyzing certification programs, and finds that pragmatic and moral legitimacy are more commonly granted to forest certification programs than cognitive legitimacy, even though the latter may be the most durable form of legitimacy.[43]

Stakeholders may evaluate these forms of legitimacy in terms of either an organization's actions and "outputs" or its essence and "inputs."[44] *Output legitimacy*, or "rule effectiveness," is the extent to which initiatives "effectively solve the issues that they target," and requires comprehensive coverage of the relevant actors, strong rule efficacy, and effective enforcement.[45] A comparison of two forest certification programs, the Forest Stewardship Council and the Sustainable Forestry Initiative, concluded that the former has higher output legitimacy

37 Patten (1992).
38 Vogel (2005: p. 73).
39 Suchman (1995: p. 574).
40 Uphoff (1989: p. 310).
41 Roberts (1992); Deegan (2002).
42 Suchman (1995).
43 Cashore (2002).
44 Suchman (1995); Mena and Palazzo (2012).
45 Mena and Palazzo (2012: p. 528).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

because it has certified a larger area of forest and induces more changes to the environmental dimensions of forest management.[46]

Such outputs, however, can be difficult to systematically quantify,[47] and so audiences may instead focus on the *input legitimacy* of "green" claims and whether the process by which the claims were generated is perceived as justified.[48] This form of legitimacy derives from a concern in democratic theory that "political choices should be derived, directly or indirectly, from the authentic preferences of citizens."[49] From this perspective, process matters as much or more than outcomes. Mena and Palazzo suggest that input legitimacy requires stakeholder inclusion, procedural fairness of deliberations, promotion of a consensual orientation, and transparency of an organization's structures and processes.[50] In essence, both who is involved and how they are involved are relevant to determining the input legitimacy of an organization or initiative.

The different signals of credibility discussed above can help either firms or evaluation organizations earn these various types of legitimacy. For example, firms and evaluation organizations can gain pragmatic and moral legitimacy if stakeholders view them as trustworthy, competent, and dynamic enough to deliver information that is relevant to either themselves or society at large. Likewise, they can gain cognitive legitimacy if stakeholders sense their traits of trustworthiness, competence, and dynamism are culturally appropriate and perceived as "predictable, meaningful, and inviting."[51] These traits generally help agents earn input legitimacy; trustworthiness, expertise, and dynamism are all traits that stakeholders may value as important inputs in the process of developing sustainability information. The one exception is transparency about outcomes, which, as noted below, can also earn agents output legitimacy.

Stakeholders, however, may disagree over which of these signals of credibility are most important, and may be more willing to grant legitimacy for some traits more than others. In this case, which principals and which signals do agents prioritize in their pursuit of legitimacy? Do they focus more on signals of safety or qualification, for example? This question relates to what Koppell terms "multiple accountabilities disorder" and Romzek and Ingraham call "cross pressures of accountability."[52] In these contexts, a common pattern is for managers to

46 Mena and Palazzo (2012).
47 Rondinelli and Vastag (2000); Blackman and Naranjo (2012).
48 Mena and Palazzo (2012).
49 Scharpf (1997: p. 19) emphasis in original.
50 Mena and Palazzo (2012).
51 Suchman (1995).
52 Romzek and Ingraham (2000); Koppell (2005).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

focus on one or two of these relationships on a daily basis with the others "being in place but underutilized, if not dormant."[53] As Mattli and Büthe find in their case study of US financial accounting standards, private sector agents with delegated public authority may focus more on the interests of their private principals if those principals are internally cohesive and have preferences that are distinct from those of other interested parties.[54] This dynamic further demonstrates the importance of credible signals of accountability.

Given that accountability is a dialectical phenomenon in which the one giving account may have significant power to dictate the terms of the relationship,[55] agents may also attempt to change or align the preferences of their principals through advertising and other means. Thus agents may have significant latitude in terms of selecting and defining the "principals" and "principles" they are accountable to, especially when the broader public is not firmly committed to a particular standard of organizational behavior.

3 Three signals of credibility

Thus the firms and evaluation organizations behind product and company environmental evaluations can seek out grants of legitimacy from particular stakeholders they are willing to be accountable to by sending signals of credibility that are salient to those stakeholders. While Berlo, Lemert, and Mertz's and Suchman's credibility and legitimacy typologies provide a useful framework for understanding the interactions between these principals and agents, they are based on multi-dimensional concepts that are difficult to measure directly (Berlo, Lemert, and Mertz's survey, for example, has 83 different dimensions).[56] Several studies have identified more specific characteristics of environmental ratings and certifications and asked both consumer and expert audiences to identify those characteristics that are most important to them.

For example, Bullock conducted a conjoint analysis-based online consumer survey asking participants to identify their most preferred characteristics of eco-labels.[57] From a set of 32 attributes that included affiliations with specific types of organizations (media, corporate, non-profit, government, and academic) and

53 Romzek and Ingraham (2000).
54 Mattli and Büthe (2005).
55 Black (2008).
56 Berlo, Lemert, and Mertz (1969); Suchman (1995).
57 Bullock (2011).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

specific content areas, independence and transparency were the two most preferred characteristics of eco-labels. The inclusion of energy and climate change criteria and expertise were the third and fourth most preferred characteristics. Similarly, a survey of over 1000 sustainability professionals found that the three most important factors for this audience, in order of importance, were objectivity/credibility of the data sources, disclosure of methodology, and experience and size of the research team.[58] These three top factors map well to the dimensions of transparency (disclosure), independence (objectivity), and expertise (research team experience) identified in Bullock's consumer study.[59]

While other characteristics may also influence stakeholder perceptions of these programs, these results suggest that transparency, independence, and expertise are among the most likely to effectively serve as specific signals of credibility for these initiatives for a broad range of audiences, from sustainability experts and professionals to the public at large. These three characteristics may overlap and complement one another (e.g., experts can be independent and transparent), but nevertheless represent distinct and independent phenomena. This section of the paper reviews past research on each of these three characteristics, and explores how they relate to the accountability, credibility, and legitimacy of environmental evaluation initiatives. In particular, it discusses how these characteristics may connect to the general dimensions of credibility outlined above (trustworthiness, safety, and dynamism), different forms of legitimacy (pragmatic, moral, and cognitive), and the interests of different stakeholder groups.

3.1 Independence: signaling distance

As discussed above, trustworthiness is one of the core dimensions of credibility, and perhaps the most commonly mentioned proxy for trustworthiness in the literature on sustainability claims is the independence of the assessment organization and its lack of conflicts of interest.[60] Smith, Palazzo, and Bhattacharya argue that a claim's credibility is particularly undermined "where consumers perceive

58 Sadowski, Whitaker, and Buckingham (2010).

59 Bullock (2011).

60 Maury (2000); Banerjee and Solomon (2003); Dando and Swift (2003); Nilsson, Tunçer, and Thidell (2004); Boström (2006); Jose and Lee (2007); Costa et al. (2009); Smith, Palazzo, and Bhattacharya (2010); Starobin and Weinthal (2010).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

firm-serving motivations rather than motivations to serve the public good,"[61] and Jose and Lee find that "companies are using third party external audits to establish the credibility of their commitment to environmental management practices."[62] The underlying logic is that companies should not be the principals for independent assessments; the more objective and distant the source of an assessment is from the source of the product, the better. The assumption is that signals of credibility either sent directly by firms or by evaluation organizations associated with those firms are inherently unreliable.

Studies by government agencies, consulting firms, and research institutions have emphasized the importance of such independence and included it as a criterion in their assessments of sustainability labels and standards verification frameworks.[63] A report commissioned by the UK government, for example, includes the presence of an independent accreditation/certification body as one of its criteria in its evaluation of the robustness of the cases surveyed. Other studies of social and environmental certifications have found that between 63 and 65% of social and environmental product certifications are implemented by third-parties, non-governmental organizations, or require external audits.[64] In its updated Green Guides, the US Federal Trade Commission mentions survey data from Cone LLC that found that 80% of consumer respondents believed that "certification by third-party organizations is important in providing oversight to ensure environmental messaging by companies is accurate."[65]

Independence maps well to Berlo, Lemert, and Mertz's safety dimension and Hovland, Janis, and Kelley's trustworthiness dimension.[66] People may be more likely to trust and feel "safe" using information coming from third parties who have fewer conflicts of interest. Building on Berlo, Lemert, and Mertz's definition of safety, if the third parties are non-profit organizations, they may also be perceived as more "altruistic" and "kind."[67] If they are professional certification organizations, academic institutions, or government agencies, they may be perceived as more "fair" and "calm."

Policies promoting independent data verification or generation by third parties may also express a normative belief in the value of civil society organizations

61 Smith, Palazzo, and Bhattacharya (2010: p. 625).
62 Jose and Lee (2007: p. 318).
63 Sadowski, Whitaker, and Buckingham (2010); Sustainable Commodity Initiative (2010).
64 World Resources Institute and Big Room (2010); Raynolds, Long, and Murray (2014).
65 See <http://www.gpo.gov/fdsys/pkg/FR-2010-10-15/html/2010-25000.htm>. Accessed December 5, 2010.
66 Hovland, Janis, and Kelley (1953); Berlo, Lemert, and Mertz (1969).
67 Berlo, Lemert, and Mertz (1969).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

as advocates of the public's interests. Likewise, they also imply that the critical locus of power and accountability should be with these organizations because of their public orientations, watchdog status, and focus on social welfare. In this sense, agents emphasizing their independence may be recruiting grants of moral input legitimacy from principals who value the role these intermediary organizations play in society. Such an approach is justified by surveys that consistently find that non-governmental organizations are society's most trusted institutions, both by the public in general terms and by sustainability professionals as evaluators of corporate sustainability performance.[68]

3.2 Expertise: signaling knowledge

Expertise has also been cited as an important aspect of legitimacy,[69] which is not only an evaluation of particular decisions but also the suitability of those who make those decisions.[70] Thus the people implementing these initiatives may have varying levels of knowledge that make them more or less qualified to determine the sustainability of a particular product or company. Assurance statements for CSR reports therefore often provide "commentary from high profile experts deemed trustworthy by the public."[71] In some cases, regulatory agencies may even delegate policy-making authority to private agents because of their pre-existing specialized expertise in particularly complex and technical issue areas.[72] There is a rich literature on the subject of expertise, and it discusses the phenomenon both generally as well as in the specific context of environmental politics.[73] One important distinction that this literature reveals is the difference between expertise from academic training ("book learning") and expertise from professional experience ("learning by doing").

Expertise is one of the core dimensions of both Hovland, Janis, and Kelley's and Berlo, Lemert, and Mertz's original typologies of credibility, and is a primary reason why the public might accept an organization as legitimate in the absence of more direct evidence of output legitimacy.[74] An emphasis on the expertise

68 See <http://www.edelman.com/trust-downloads/press-release/>. Accessed January 19, 2015; Sadowski, Whitaker, and Buckingham (2010).
69 Dando and Swift (2003); Gifford (2010).
70 Uphoff (1989).
71 Dando and Swift (2003: p. 196).
72 Mattli and Büthe (2005).
73 Ericsson and Smith (1991); Chi, Glaser, and Farr (1988); Jasanoff (2003).
74 Hovland, Janis, and Kelley (1953); Berlo, Lemert, and Mertz (1969).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

behind an assessment process may represent a commitment to scientific knowledge as the best way to ensure the validity of an evaluation (and dealing with sustainability challenges more generally). From this perspective, it is the scientists and experts who should be trusted to solve society's environmental problems and evaluate claims of "greenness." Following this logic, organizations that hire experts with relevant expertise are more likely to produce valid environmental assessments.

An emphasis on expertise may also represent a normative commitment to the rigorous pursuit of truth as a fundamentally important social value. It may thus signify an attempt to activate a sense of cognitive input legitimacy; like technical evaluations in other domains, assessments of sustainability should naturally be conducted by experts with relevant technical knowledge, and to think otherwise is "unthinkable."[75] Such a dynamic would explain why over 1000 sustainability experts rated the experience and size of the research team as one of the three most important factors in determining the credibility of a corporate sustainability rating.[76]

3.3 Transparency: signaling openness

Transparency is one of the most frequently cited dimensions of legitimacy in the peer-reviewed literature on eco-labels and ratings,[77] and encompasses the more specific concepts of traceability and auditability.[78] Hess, for example, points out that "to have meaningful stakeholder engagement requires that we first have a robust information-based transparency policy with comparable data,"[79] while Fung asserts that democratic transparency requires the disclosure of rich, usable, and actionable information whose availability is proportionate to the risks to which it is relevant.[80] Auld and Gulbrandsen differentiate between *procedural transparency*, which refers to the "openness of governance processes" and relates to the concept of input legitimacy, and *outcome transparency*, which "deals with the substantive ends of a given policy intervention" and can contribute to output legitimacy.[81]

75 Suchman (1995).
76 Sadowski, Whitaker, and Buckingham (2010).
77 Dando and Swift (2003); Nilsson, Tunçer, and Thidell (2004); Fung, Graham, and Weil (2007); Hess (2007); Auld and Gulbrandsen (2010); Mena and Palazzo (2012); *ibid*.
78 Nilsson, Tunçer, and Thidell (2004); Boström (2006).
79 Hess (2007: p. 471).
80 Fung (2013).
81 Auld and Gulbrandsen (2010).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

A survey by World Resources Institute and Big Room found that 87% of the programs that responded to the survey make their certification criteria public, while 44% have measured the environmental and social impacts of their labels.[82] Such transparency likely taps into all three of Berlo, Lemert, and Mertz's dimensions of credibility – more transparent programs may evoke not only a sense of safety and qualification but also dynamism.[83] By frankly and confidently revealing the methods behind their ratings, transparent evaluations convey a sense of "boldness, frankness, and energy," the core characteristics of a dynamic organization.[84] Transparency also requires an investment of time and resources, and is likely to be less onerous and risky for organizations that have the qualifications to conduct valid analyses. Transparency may also elicit a sense of trustworthiness – an initiative that is open about itself reduces concerns about ulterior motives and hidden conflicts of interest.

While transparent programs do not necessarily allow external stakeholders to directly participate in their decision-making processes, they nevertheless convey a commitment to the value of openness and the ability of the public to reward excellent programs and punish fraudulent ones. They also place an implicit normative value on information as a mechanism to inform, educate and mobilize the public about important social problems. A commitment to transparency suggests that it is the public to whom these programs should be accountable to – not only experts or organizations claiming to represent society's interests.

Agents that are highly transparent may therefore be seeking grants of pragmatic input legitimacy from the public by implicitly signaling a respect for their ability to evaluate and utilize the provided information based on their own values. To the extent that their openness is itself a public good that encourages transparency and improved sustainability performance, transparent firms and evaluation organizations may also be eliciting grants of moral input legitimacy. If they are transparent about their sustainability outcomes, principals may also perceive them as having high moral and pragmatic output legitimacy. And if they reflect a broader norm of expected organizational behavior (that environmental evaluation initiatives should be transparent about themselves, especially given that they often demand transparency of the firms they are evaluating), transparent agents may also be eliciting a sense of cognitive legitimacy as well.

82 World Resources Institute and Big Room (2010).
83 Berlo, Lemert, and Mertz (1969).
84 *Ibid*.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

3.4 Connecting credibility signals, moral principals, and forms of legitimacy

Transparency, independence, and expertise may therefore all serve as useful signals of the legitimacy of competing environmental evaluation initiatives, which may be produced by either firms or evaluation organizations. As Figure 2 illustrates, each emphasizes different aspects of credibility, suggests accountability relationships to different principals, and may elicit different forms of legitimacy. By detecting the relative prevalence of these different signals among existing environmental evaluations of products and companies, we can determine to whom these programs are most likely signaling their credibility and whom they most likely identify as their primary principals. As suggested above, expertise signals suggest accountability to experts, independence signals suggests accountability to advocates, and transparency suggests accountability to the broader public, which includes but is not limited to experts and civil society representatives.

While evaluation initiatives may have additional accountability relationships to other principals (e.g., government, investors), this framework focuses specifically on non-profit civil society organizations and experts as principals because of their high levels of perceived trustworthiness (in the case of the former) and perceived competency (in the case of the latter). If evaluation initiatives can earn grants of legitimacy from these principals by demonstrating their independence or expertise, then other principals may follow suit. Alternatively, they can seek grants of legitimacy more directly from a broader range of principals – including not only experts and advocates but also consumers, policymakers, and other public

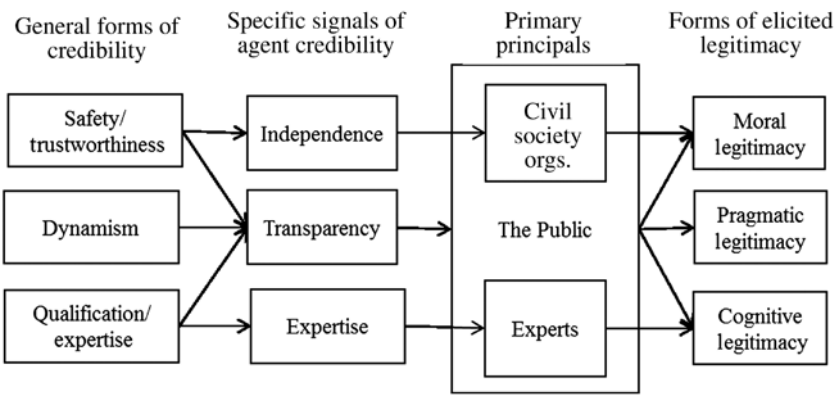

Figure 2: Credibility-accountability-legitimacy conceptual framework.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

stakeholders – by sending signals of transparency. Given that this conceptual framework suggests that transparency appeals to the broadest audience, incorporates multiple forms of credibility, and elicits multiple forms of legitimacy, a reasonable hypothesis is that it is the most common signal being sent by these programs.

3.5 Detecting the signals

The literature on eco-labels and ratings cited above supports this paper's focus on independence, transparency, and expertise as key signals of credibility, and several studies provide an initial indication of the prevalence of these different signals. Many also effectively document the procedural and substantive characteristics of the programs they survey – in some cases, better than the programs do themselves. The level of technical detail collected by these reports is often outstanding, and enables robust comparison of the methodologies of the cases covered. Taken as a group, however, they also have several important shortcomings that limit the usefulness of their data and analyses.

For example, many tend to focus on the methodological elements (i.e., the actual methods used) of these programs at the expense of their more procedural aspects (e.g., transparency, independence, and expertise). Those that do cover these latter attributes do not specify them in much detail and generally ignore their underlying complexity. They therefore tend to be either too detailed, focusing on technical details that most audiences do not have the capacity to evaluate, or too general, avoiding the multiple dimensions of key attributes that most audiences are more likely to notice and assess.

In terms of the scope of these studies, they primarily focus on product evaluations and seldom include evaluations of corporate environmental performance – only two of the 13 major initiatives reviewed above cover such corporate evaluations, despite the fact that such initiatives have become increasingly prominent in the US marketplace (e.g., Newsweek's Greenest Large Companies Ranking, Fortune Magazine's Green Giants List). Assessments of companies and products share similar purposes and functions, and should be included together in analyses of sustainability claims.

Secondly, most of these studies are global in scope, surveying eco-labels and green ratings available around the world. This sampling frame is useful in understanding how the characteristics of this form of governance may vary across national boundaries, but risks only including the most well-known initiatives in different countries and excluding less well-known labels and ratings that nevertheless may be important to particular stakeholder groups. A notable exception

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

is Raynolds, Long and Murray's study, which includes a large sample of certification systems that operate in the US.[85] However, the sample excludes cases that are implemented by the government or individual firms and cases that do not have oversight over or specific requirements for firms, such as ISO 14001. While providing interesting data on a range of different characteristics, their dataset also does not provide any data on signals of expertise or transparency.

Many of these initiatives have important methodological shortcomings as well. Several, for example, solicit information from the organizations behind these labels through surveys or questionnaires, as opposed to only using publicly-available information provided by a randomly selected set of cases.[86] Such a process, while enabling the collection of more detailed information from a subset of cases willing to provide it, risks significant bias in the sampling process and increases the probability that the results do not reflect the actual distribution of characteristics across the full universe of existing and relevant initiatives. A comprehensive analysis of the phenomenon of green claims should include not only the most popular and responsive of programs, but the full range of programs that are available to the public. Several of the research reports cited also do not clearly describe their sampling process, and so their data may be non-random and biased as well. Given these methodological issues, none of the existing studies provides a reliable analysis of the signals of credibility being produced by environmental evaluation initiatives.

4 Methods

In order to address this gap in the literature, a study was designed to avoid these shortcomings and provide a more accurate analysis of these signals. The goal of this study was to rigorously collect data on the organizational, methodological and procedural characteristics of a broad range of not only product eco-labels but also sustainability ratings, rankings, awards, boycotts, databases and reviews related to both products and companies. These are all mechanisms that principals can use to monitor their moral agents, and focusing only on one form does not capture the full range of information available to them. Rather than collecting data that is either too general or too detailed, data was collected at the level of detail that is most likely serving as signals of credibility to the public. Instead of

85 Raynolds, Long, and Murray (2014).

86 Sadowski, Whitaker, and Buckingham (2010); World Resources Institute and Big Room (2010); Raynolds, Long, and Murray (2014).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

relying on non-random responses to questionnaires, data was collected directly from the websites of a broad range of 245 initiatives. The text from these websites was systematically downloaded, coded, and analyzed using methods that are explained in detail below. This section describes the methods used to create such a dataset, including the sampling process, data collection protocols, data quality assurance procedures, and data analysis process.

In order to enable an in-depth analysis of the credibility signals that are being sent to American consumers from a broad range of initiatives, including not only more well-known but also less well-known programs, the sampling frame was limited to initiatives that are relevant to the US market. The US has the world's largest economy and the largest number of relevant eco-labels,[87] making it an important case to examine. While a comparison of signaling patterns across different types of economies might yield interesting insights and different results, this paper provides an in-depth and relatively comprehensive analysis of those signals within one particular economy where they are particularly prevalent. As Raynolds, Long and Murray also note, national laws, market characteristics and social movement pressures shape the need for private regulations, and further justify an empirical focus on initiatives available within a specific country.[88]

The study's sample was selected through a multi-step process that first involved aggregating several online databases of environmental evaluations and lists of relevant programs, including Ecolabelling.org, Ecolabels.org, AllGreenRatings.com, the Global Ecolabelling Network, and ISEAL. The initial sample also included initiatives identified in news reports, academic articles, blogs, and similar sources of information between 2006 and 2008. Initiatives identified through a series of systematic keyword searches on Google for "eco-labels," "green ratings," and other keywords across a set of 10 product categories (e.g., electronics, toys, etc.) were also included. This process resulted in a list of 471 initiatives, identified through the end of 2008 (thus programs introduced after 2008 are not included). They include programs developed by government agencies, non-profit organizations, academic institutions, for-profit enterprises, and media outlets.

In order to ensure that the sample of cases is an accurate and unbiased sample of information-based environmental initiatives, programs were excluded that did not meet the study's sampling frame, which limits the included cases to *"information-based environmental governance initiatives that generate publicly-available environmental evaluations of products or companies that make products*

87 The Ecolabel Index (ecolabelindex.com) lists 201 eco-labels found in the US, nearly twice the number found in the next two countries (Canada and Germany) combined.
88 Raynolds, Long, and Murray (2014).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

that are generally available in the US marketplace." In this context, a governance initiative is an intentional, planned effort to exert power over others to encourage collective action and create public goods. This definition excludes both internally-oriented information-based environmental management programs that are not made public as well as anonymous, hearsay, and generic claims, such as "natural" or "recyclable," that do not have a single, traceable source. It also does not include corporate sustainability reports, which are generally descriptive and not evaluative forms of information.[89]

Excluding duplicates, initiatives that did not meet this sampling frame, and initiatives that overlap, replicate, or are part of a broader program reduced the final sample size to 245 cases. The cases most commonly focus solely on evaluating the performance of products (47% of cases), followed by companies (27%) and facilities (15%). The remaining 20% evaluate some combination of product, company, and facility performance. The top five sectors covered by the sample are food (44 cases), household products (39), apparel (38), personal care (36), and electronics (35). Most of the cases (94%) are implemented not by the firms being evaluated but by different types of evaluation organizations, including professional certifiers, advocacy groups, online shopping sites, industry associations, academic institutions, government agencies, media outlets, and multi-stakeholder coalitions.[90] Table 1 provides summary statistics about the types of initiatives in the sample. Certifications are the most common form of evaluation found in the dataset (41% of cases), while over a third of the cases are information "hybrids" that provide multiple forms of evaluation (e.g., both a rating and a ranking).

Data about this sample of 245 cases were then collected through a rigorous and comprehensive content analysis process. Content analysis is a "research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use,"[91] and has been used extensively to analyze corporate annual reports and sustainability disclosures.[92] While utilizing

89 This distinction follows Kuklick's (1969: p. 92) definitions of descriptive meaning as "the meaning of words…which describe or state facts" and evaluative meaning as "the meaning of words…which are closely connected with choice, decision, and action" and includes "emotive," "laudatory," "commendatory," "prescriptive," and "normative" meaning.

90 This statistic is somewhat misleading, however, because as Bullock (2015) shows, firms can influence these initiatives in many ways beyond directly leading and implementing them. Nevertheless, it does reveal the proliferation and prominence of sustainability evaluation organizations that are at least on the surface independent of the firms being evaluated.

91 Krippendorff (2012: p. 24).

92 Gray, Kouhy, and Lavers (1995); Buhr (1998); Beattie, McInnes, and Fearnley (2004); Milne, Walton, and Tregidga (2009).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

Table 1: Types of initiatives in dataset (See Online Appendix Table A1 for definitions of types).

Type	Stand-alone	Within hybrids	Total
Certification	78	23	101
Award	30	43	73
Rating	24	32	56
Database	8	42	50
Rated Certification	20	10	30
Review	1	24	25
Boycott	2	13	15
Ranking	0	11	11
Hybrids (More than one type)	82	–	82
Total	245	–	–

a similar method, this study instead focuses on analyzing the content provided in environmental evaluations of products and companies. It uses a research design incorporating both mechanistic and interpretative elements in order to both identify the presence or absence of particular themes and interpret the meaning of those themes.[93] This process involved identifying a set of characteristics relevant to these themes, defining these characteristics using a set of variables or "codes," and using these codes to analyze the text from the websites of the sample cases. While some initiatives communicate with their audiences through other media as well, the internet is the primary means by which all of these programs provide comprehensive information about themselves to the public and is therefore the most suitable source of data for this study.

To develop the initial set of codes, an analytic inductive process building on insights from the literature was used to code the website text from an initial set of 40 cases.[94] Text segments were coded by theme and not limited to individual words, sentences, or paragraphs, allowing the coding process to capture signals at multiple levels of textual resolution.[95] Similar to the two-step approach used by Beck, Campbell, and Shrives,[96] four other coders used this same set of codes to code the text of the same 40 cases, and discrepancies in the results revealed that transparency, independence, and expertise have more dimensions than previous research has identified. This process allowed for a more nuanced interpretation

93 Beck, Campbell, and Shrives (2010).
94 Bansal and Roth (2000).
95 Beck, Campbell, and Shrives (2010).
96 *Ibid*.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

of the meaning associated with these themes and a more granular definition of their associated codes. The multiple dimensions of these characteristics, which are described in more detail below, were incorporated into the coding system used to analyze the full sample of cases.

This coding system was documented in a detailed codebook, which includes 223 binary codes indicating the presence of particular characteristics relating to the methods, content, and organizational backgrounds of the sample cases. Two coders used this codebook and the qualitative coding software MaxQDA to manually code 2535 webpages and PDF documents downloaded from the websites of the 245 cases in the sample. This process was completed between April 2009 and September 2010, and produced a dataset of 9829 coded text segments. This paper reports on the 40 binary codes and 2245 coded text segments documenting the transparency, independence, and expertise of these cases. The definitions of these 40 codes are provided in Online Appendix Table A1, and examples of each are provided in the sections below. Because the websites of these initiatives had such extensive amounts of text that coders had to manually download and code, programs introduced after 2008 and website text updated after September 2010 are not included in the study.

The analysis process involved compiling the data produced from this coding analysis, testing the inter-rater reliability of this data, and checking it for errors and inconsistencies. In order to ensure the replicability and inter-rater reliability of the data, both coders coded a random sample of 25 cases, or approximately 10% of the overall sample, and compared the results for discordances. The average level of agreement for the coded data is 91%, with an average Kappa score of 0.28, indicating a fair level of agreement.[97] While some of the codes have lower Kappa scores, they are likely due to the lower overall prevalence of the characteristics they were measuring. Those with significantly lower levels of agreement (Kappa values<0) are mentioned in the relevant sections below.

5 Results

5.1 Signals of independence

The inductive analytic process described above revealed that independence has several dimensions that are often overlooked by analysts of these initiatives. The first of these dimensions is the *type of independence* – has the data been generated by independent organizations or only verified by such organizations?

97 Landis and Koch (1977).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

Independent generation implies full control of the data from collection to analysis to delivery, while verification indicates external monitoring of a self-assessment process that has higher potential for fraud. A second important dimension is the *source of the independence* – is the data generation or verification performed by the evaluation organization itself, or is it conducted by an organization that has been accredited or contracted by the evaluation organization? Increasingly, "third party" certification systems are assigning the strategic roles of standard-setting and administration and the operational roles of monitoring and assessment to separate organizations. A third dimension is the *level of independence* – is all of the data independently generated or verified, or only some of it?

Three sample text segments provide examples of each of these different characteristics and demonstrate how they were coded. The website of the Environmental Protection Agency's WaterSense program states, "All products bearing the WaterSense label must be tested and certified by an approved third party laboratory to ensure they meet EPA water efficiency and performance criteria."[98] This is an example of a text segment that was coded as *all data* ("all products…must be tested"), *independent generation* (external labs, not EPA, are conducting the tests), and *contracted/accredited organization* ("an approved third party laboratory"). As a second example, the website of B-Corp states, "When a company becomes Certified they must submit documentation for approximately 20% of their answers to the B Survey…10% of B Corporations are audited every year… [by B Lab auditors]." This text segment was coded as *some data* (only 10% are audited), *independent verification* (data is submitted by the company), and *evaluation organization* (B Lab auditors). In cases where the source or type of independence was unclear, such as the phrase "third-party, independent validation and verification" found on Rainforest Alliance's website, the text segment was coded as *evaluation organization* and *independent verification* by default.

Almost 40% of the cases in the sample verified or generated at least some of their data. Slightly over 14% of the cases generated their own data independently of the organizations being evaluated, and slightly over 33% had mechanisms in place to verify the accuracy of the data they received from the organizations they were evaluating. Almost 30% of the cases verified or generated all of their data, and nearly 10% verified or generated some of their data. Approximately 28% of the cases have other organizations generate or verify their data, while just under 18% generate or verify their information themselves. Figure 3 presents a more granular view of these data. The proportion of cases that use independently verified or generated data was not significantly different for cases implemented by

98 "All products" refers to all product types – a sample of products from each product type is selected for testing by the laboratory.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

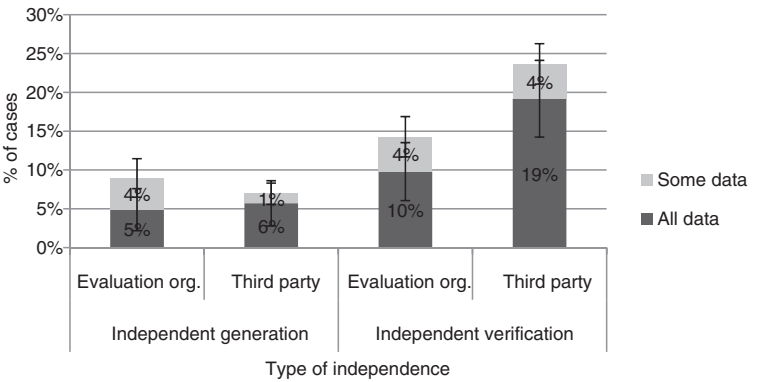

Figure 3: Types and levels of data independence (By % of cases).
Error bars indicate 95% confidence intervals for both the "some data" and "all data" sample proportions. Because of overlap between categories, the statistics presented in the figure do not always add up to the percentages mentioned in the text.

firms than for cases implemented by evaluation organizations (one-sided Fisher's exact test=0.575).[99]

An additional dimension of independence is the type of peer review, if any, that is used in the evaluation process. Both the methods used in the evaluation and the data collected can be peer reviewed, and the review can be conducted by individuals with varying levels of expertise who work inside or outside the firm or evaluation organization. An example of *data peer review* comes from the Rainforest Alliance, which states that a team of trained specialists writes an assessment report of a farm or forest that has applied for certification, and this report is then "evaluated by an independent, voluntary committee of outside experts (i.e., peer reviewed)." An example of *method peer review* comes from Protected Harvest, which states that its "standards are peer-reviewed by the scientific community and then must be approved by the distinguished environmentalists on the Protected Harvest board." Approximately 5% of programs mention peer review processes for their methods, and 4% mention peer review processes for their data. Less than 2% of the cases specified the expertise of the individuals conducting the peer review process. For example, one text segment states that "BASF's eco-efficiency was carefully examined and evaluated by David R. Shonnard, PhD, an independent expert in green engineering," and goes on to describe his academic credentials.

99 Cases implemented by firms include any case in which the organization conducting the evaluation is evaluating its own products or performance. This includes manufacturers as well as retailers who evaluate their own branded products and products of other companies.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

5.2 Signals of expertise

A second important signal of credibility is the expertise of the individuals who conduct the assessment. As mentioned above, expertise can be produced through academic training or from professional experience. Academic training can be further categorized as general training or training that is directly relevant to the organization's work. In order to capture these dimensions of expertise, text segments were coded as general academic training, relevant academic training, and relevant professional experience. As an example of *general academic training*, the CarbonNeutral website states that its Executive Vice President "holds an MBA with Distinction from the Stern School of Business at NYU and a Bachelors degree in Psychology from UCLA." The Bird Friendly Coffee website states that the Director of the organization behind the certification has a PhD in ornithology from the University of California, Berkeley, which is an example of *relevant academic training*. The website of the 100 Best Corporate Citizens provides an example of *relevant professional experience*; its Director of Research is described as having "more than a dozen years of experience supporting institutional investors with research and software tools for values-based investing and proxy voting."

The coding data indicates that nearly 1 out of 5 cases (18%) claim that at least one staff member working on the initiative has relevant professional background and expertise (i.e., substantive, full-time past work on environmental or social issues). Slightly over 10% claim to have staff with academic training (masters or above) that is relevant to environmental or social issues, while approximately 7% claim to have staff with academic training (masters or above) that does not have a clear relationship to the work of the initiative (see Figure 4). While approximately 25% of the cases implemented by evaluation organizations make at least one of these claims of expertise, none of the initiatives implemented by firms make any

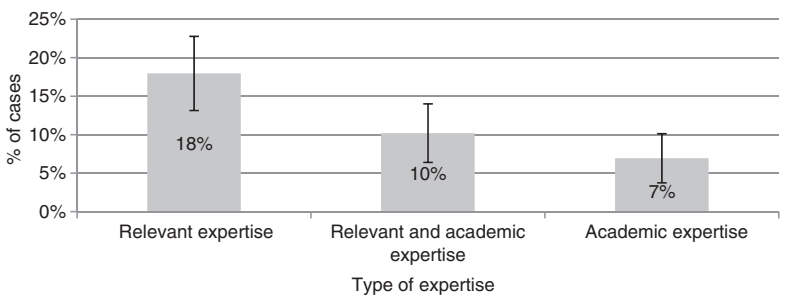

Figure 4: Types of expertise mentioned (By % of cases).
Error bars indicate 95% confidence intervals for each sample proportion.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

claims of expertise. Firms are therefore significantly less likely to signal their expertise than evaluation organizations (Fisher's one sided exact test=0.015).

It should be noted that the inter-rater reliability for two of these expertise-related attributes was among the lowest in the broader set of data collected. The probability that the agreement between the raters was due to chance is between 60 and 66% for the academic and relevant academic expertise codes, and their Kappa scores were both <0 (−0.04 and −0.07, respectively). This may be more due to the low prevalence of these characteristics (the calculated Prevalence Index was 0.88 and 0.84, respectively, for these two criteria), and less due to the reliability of the coding process.[100] However, expertise may indeed be difficult for both coders and more general audiences to recognize and agree on, especially given the many different forms it can take.

5.3 Signals of transparency

Similar to the cases of independence and expertise, the process of developing the codes for transparency revealed that this concept is more complex and multi-dimensional than is usually acknowledged. The analytic inductive process described above identified seven different ways that initiatives can be transparent about their evaluation processes. The coding process also distinguished between "limited" and "strong" statements of transparency. Strong criteria transparency, for example, indicates that all of the criteria are fully explained, with at least a sentence about what is being measured and what data is being used for each cited criteria, while limited criteria transparency indicates that some but not necessarily all the criteria are listed, and they may or may not be described in any detail. Examples of each type of transparency are provided in Online Appendix Table A2, and the coding results are discussed below.

Figures 5 and 6 summarize the data for each dimension of transparency. *Criteria transparency* refers to the extent to which a case describes the criteria they use in their evaluation of either products or companies. Sixty-seven of the cases, or 27%, describe some but not all of their criteria (*limited criteria transparency*), while 51% describe their criteria in full detail (*strong criteria transparency*). *Data transparency* refers to whether an initiative provides the actual data underlying the evaluation on their website. The content analysis of the full sample indicates that over 40% of the cases provide none of their underlying data, 35% provide some but not all of their data (*limited data transparency*), and 24% provide all of their underlying data (*strong data transparency*). *Source transparency* refers

100 Sim and Wright (2005).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

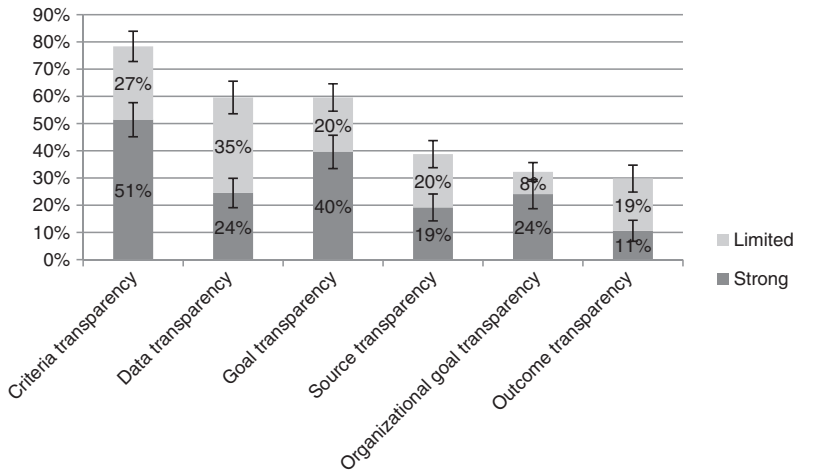

Figure 5: Types and levels of transparency (By % of cases).
Error bars indicate 95% confidence intervals for each "limited" and "strong" sample proportion.

to whether a case provides a list of the sources of the data that is the basis of its evaluation. Approximately one fifth of the cases have *limited source transparency* (some but not all of the sources are listed), and another fifth have *strong source*

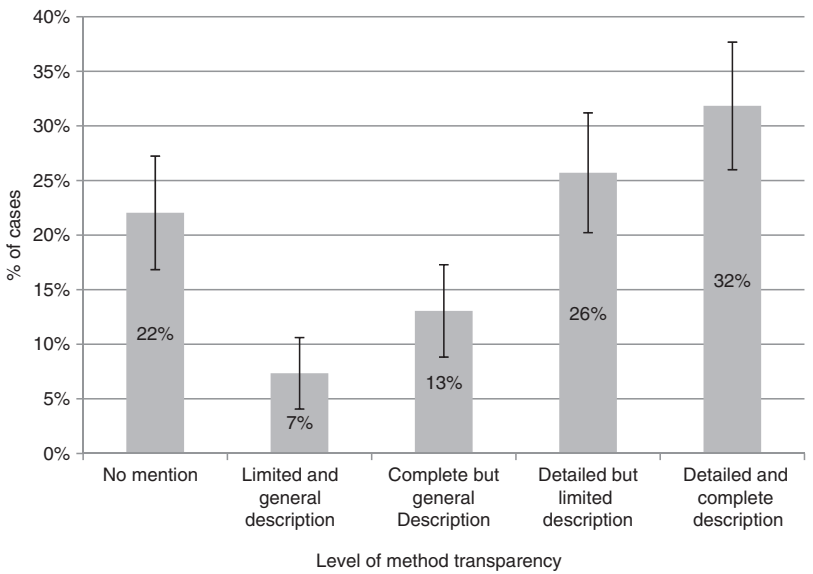

Figure 6: Levels of method transparency (By % of cases).
Error bars indicate 95% confidence intervals for each sample proportion.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

transparency (all of the data sources are listed). Three fifths do not provide any information about their data sources.

Two other aspects of transparency are the extent to which these programs are open about their goals (*goal transparency*) and the goals of the organizations behind them (*organizational goal transparency*). The primary distinction between strong and limited goal transparency is whether the goals are stated explicitly on the website or if they are provided more implicitly in the text without specifically mentioning the words such as "goal," "mission," or "purpose." Approximately 20% have such limited goal transparency, nearly 40% of the cases have strong goal transparency, and over 40% do not mention their goals at all. Approximately one third of the cases state the goals of the organization behind the initiative (either explicitly or implicitly), while two-thirds do not. Overall, 25% of the cases do not provide any information about the goals of either the initiative or the organization behind it.

The objectives of these cases may also be apparent through another form of transparency – providing information about the environmental outcomes produced by the initiative (*outcome transparency*). Strong outcome transparency requires specific claims about the actual benefits that are produced by the initiative, while limited outcome transparency includes general claims regarding the potential social or environmental benefits of an initiative. Slightly over 10% make specific claims regarding the actual benefits created by the program, while nearly 20% make general claims about the potential social or environmental benefits of the initiative but do not discuss actual outcomes. Over 70% do not mention either real or potential outcomes of their program. This form of transparency is closely related to the output legitimacy of these initiatives, but does not assess the quality of their outputs, only whether they are discussed.

Method transparency refers to the level of detail provided about how the evaluation was conducted. Given the complexity of this characteristic, four binary codes indicating increasing levels of method transparency were used to document this characteristic (see Figure 2). These codes captured two dimensions of method transparency – the *specificity* of the information (detailed versus general) provided about the methods used and the *completeness* of that information (complete versus incomplete). Approximately one-third of the programs provide a *detailed and complete description* all the methods, algorithms, and processes necessary to replicate the results of their assessment, 26% provide most but not all of the information necessary to replicate their results (*detailed but limited description*), 13% provide a *complete but general description* about their evaluation process, and 7% provide a *limited and general description* of their methods. The remaining 22% provide no information on their methods at all. Cases implemented by firms were significantly less likely than cases implemented by evaluation organizations

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

to provide detailed and complete methodological descriptions (one-sided Fisher's exact test=0.003) and more likely to provide limited and general descriptions (one-sided Fisher's exact test=0.087). No significant difference was found between cases implemented by firms and cases implemented by evaluation organizations for the other forms of transparency.

6 Discussion

These results clearly show that the environmental evaluation initiatives in this sample are more commonly signaling their credibility with transparency than with independence and expertise.

While the levels of specific types of each attribute vary considerably, the most common form of transparency – criteria transparency – was much more commonly found (in 78% of cases) than the any form of independence (40% of cases) or expertise (25% of cases). This is true for cases implemented by both firms and evaluation organizations. Furthermore, nearly 98% of the cases provide at least limited transparency – and 85% provide strong transparency – on at least one of the seven dimensions of transparency investigated. Following the accountability-based conceptual framework presented above, these results suggest that these agents are seeking legitimacy primarily from the general public and principals who value transparency. Principals may value transparency because it taps into multiple forms of credibility – safety, qualification, and dynamism, and agents may prioritize it because it has the potential to elicit grants of several types of legitimacy – pragmatic, moral, and cognitive – from these audiences.[101]

As discussed above, transparency may also represent an organization's sense of accountability to the public at large, as opposed to specific stakeholder groups. In contrast, independence, the second most common signal of credibility detected in the sample, may represent accountability to civil society organizations that are committed to improving social welfare. The nearly 40% of the sample that mention the independence of their data may be signaling their trustworthiness to principals who want to support sustainable companies but do not trust firms to assess their own sustainability. Believing that they are independent from the self-interested motives of firms, these principals may thus provide grants of input and moral legitimacy to these agents.

The least common signal of credibility was expertise, which was mentioned by <25% of the cases. Signaling qualifications, expressing accountability to

101 Berlo, Lemert, and Mertz (1969); Suchman (1995).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

experts, and eliciting grants of cognitive legitimacy (based on a general norm that experts should conduct technical assessments) appear to be much lower priorities for the agents represented in this sample. This appears to be particularly true for cases implemented by firms, which did not provide any expertise signals in the sample. These cases, while providing similar overall levels of transparency, were also less likely to be fully transparent about their methods, which is the form of transparency most directly related to expertise. This further suggests that experts are not the primary audience of the firms behind these initiatives.

6.1 Managerial implications

This research has important implications for managers at both firms and evaluation organizations. As the literature on sustainability disclosures has shown, firm managers have many possible motivations for disclosing information about their social and environmental performance.[102] Research showing over-reporting of positive news and under-reporting of negative information has raised concerns about the credibility of this self-disclosed information.[103] The small number of firm-implemented initiatives in the sample suggests that many firms are opting not to evaluate their own performance. Managers may instead be utilizing external organization to rate or certify the sustainability of their products or operations in order to enhance the credibility of their claims.[104]

As the data shows, initiatives implemented by both firms and evaluation organizations can and do send similar signals of credibility, but their principle audiences (and potential principals) may perceive them differently. A claim of independence coming from a manufacturer may be much less credible, for example, than a similar claim from a non-profit certification agency.[105] Firm manager therefore must balance this potential loss of credibility with the benefits of greater control associated with self-evaluations. If they choose to conduct their own performance evaluation, what signals should they send about it? Or, if they opt for an external evaluation organization, which one(s) should they use and endorse?

102 Deegan (2002); Paulraj (2009).

103 Adams (2004); Cho et al. (2012); Boiral (2013).

104 Hickle (2007); Boiral and Gendron (2011); Parguel, Benoît-Moreau, and Larceneux (2011).

105 Recent research has shown that consumers prefer green product claims that have been verified by an independent third party more than first party claims made by companies [UL Environment and Shelton Group (2014)].

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

This paper provides a useful framework that can help managers address these parallel questions. Ideally, firms should implement or select programs that send signals of trustworthiness, expertise and independence and have high levels of pragmatic, cognitive, and moral legitimacy. However, few programs have all of these attributes; only two cases (Rainforest Alliance and Bird-Friendly Coffee) in this dataset were coded for the strongest signals across the three signal categories – relevant academic expertise, strong method and outcome transparency, and independently-generated data. Only two additional cases, Best Aquaculture Practices Certification and GREENGUARD, were coded for relevant academic expertise and independently-generated data, and only three additional cases, the Marine Stewardship Council, the Greener Electronics Guide, and the WaterSense Label, were coded for strong method and outcome transparency and independently-generated data.

Correlations across the 34 different signals discussed in this paper are also low; the average positive correlation is 0.09, the average negative correlation is −0.04, and the highest correlation between signals is 0.26 (between "detailed and complete method transparency" and "all data independently generated by a third party"). Correlations across aggregations of these variables are also low – the correlation between cases with any strong transparency measures and any independence measures is 0.10, between cases with any strong transparency measures and any expertise measures is 0.16, and between cases with any expertise measures and any strong transparency measures is 0.24.

Thus a lack of overlapping signals may demand tough choices about priorities among these dimensions. A firm's selection of an evaluation mechanism may therefore depend on who the firm prioritizes as its "moral principals." Given that potential principals have different signal preferences, this is a classic example of Romzek and Ingraham's "cross pressures of accountability."[106] Which stakeholders does it want to appear credible and accountable to for its sustainability practices? Which sources and types of legitimacy are its highest priorities? These questions are relevant to evaluation organization managers as well. What signals of credibility should they send to attract both clients (e.g., firms) and users (e.g., consumers)?

These questions require a thoughtful assessment of the different elements of the conceptual framework presented in this paper, from the general dimensions of credibility and legitimacy to the specific characteristics of independence, expertise, and transparency. The framework suggests that if firm and evaluation organization managers prioritize the credibility and cognitive legitimacy associated

106 Romzek and Ingraham (2000).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

with expertise and view themselves as most accountable to experts, then they should emphasize signals of expertise. If they prefer to enhance their trustworthiness and moral legitimacy and build support among civil society organizations, then they should emphasize signals of independence. If they want to attract the support of the general public and recruit a mix of pragmatic, moral, and cognitive legitimacy, then they should be as transparent as possible. The data further suggest that transparency should also be their first priority if they want to adhere to the most common norm of credible behavior, while expertise should be a high priority if they want to send a less common signal of credibility that may enable them to "stick out from the crowd."

6.2 Societal implications

This research also raises important questions for the stakeholder groups – consumers, employees, communities, activists, policymakers, and procurers – who must decide if they will use these environmental evaluations to monitor and evaluate the social and environmental performance of firms. This paper reveals the complex and multi-dimensional nature of the credibility signals that these programs are sending to stakeholders, and highlights the difficulty of using them to differentiate between programs. Which type of independence is preferable? Which source of expertise is the most important? How should the different forms of transparency be weighted? These different characteristics may have limited added value as signals of credibility if stakeholders do not have further guidance and clarity about their meaning and importance. Just as voluntary standards of disclosure have been developed for corporate sustainability performance,[107] general standards and clear definitions of transparency, independence, and expertise could be developed that take into account this complexity and incorporate it into a straightforward checklist, rating, or certification. Such standards could build on and further refine the definitions provided in this paper, and might improve the quality of these signals and their usefulness to stakeholders.

For example, clear standards for different types of independent data verification and generation – including those that require civil, discursive and/or consensual engagement with stakeholders[108] – could be developed. Such standards could clarify the confusion around the term "third party," and create clear categories of independence that differentiate between firm-paid versus

107 Gilbert and Rasche (2008); Farneti and Guthrie (2009); Fifka and Drabble (2012).
108 Baur and Palazzo (2011).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

non-firm-paid, governmental versus non-profit monitoring, and standard-setting versus standard-checking.[109] This would help address Starobin and Weinthal's concern about the current black-box nature of many "third-party" certifiers.[110] Standards for relevant professional and academic expertise would be useful as well, particularly given the relatively low prevalence of expertise signals in the sample. The US Green Building Council's LEED-related credentials for architects is one example of such a standard, and could be replicated for professionals involved in the design and operation of eco-labels and other forms of environmental evaluation.[111] Such a standard would enable easier comparison among expertise claims.

Likewise, the development of a "transparency index" that takes into account the seven different forms of transparency outlined in this paper would assist stakeholders in evaluating how broadly transparent these programs are. Polls of both experts and the public could be used to determine weights for the different types of transparency. Given that outcome transparency is the only type that is related to output legitimacy, stakeholders may choose to weight it more highly.

Given the overall low prevalence of many of the signals of credibility found in the sample, some stakeholders may be motivated to require mandatory disclosures from firms and evaluation organizations about their independence, methods, and expertise. Several scholars support the mandatory disclosure of sustainability information from companies,[112] and research has shown that such disclosure laws can increase the amount of information available about firm environmental performance.[113] Nevertheless, disclosure may still vary considerably across firms because of technical loopholes, limited enforcement, and different interpretations of the law,[114] and may be motivated more by concerns about stakeholder opinion than government mandates.[115] Mandatory disclosure laws can also be particularly challenging to pass,[116] and can undermine the value of the credibility signals being sent. Signaling theory suggests that if all firms are required to send similar signals of expertise, for example, it becomes more difficult for

109 Some have described a "third party" as an external group separate from manufacturers, industry associations, and governmental bodies [Gereffi, Garcia-Johnson, and Sasser (2001)] while others have described it as "external" or "outside" but still paid for by the company [Prakash and Potoski (2006: p. 22); Starobin and Weinthal (2010)].

110 Starobin and Weinthal (2010).

111 See <http://www.usgbc.org/leed/credentials>. Accessed April 27, 2014.

112 Williams (1999); Adams (2004).

113 Frost (2007); Fung, Graham, and Weil (2007).

114 Graham (2002: p. 142–143); Frost (2007).

115 Wood and Ross (2006).

116 Hamilton (2005).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

principals to use expertise to differentiate between agents.[117] An alternative approach would be to create a virtual "information marketplace" that enables direct comparisons of credibility signals across ratings and labels. Coupled with the voluntary standards and indices of independence, expertise, and transparency discussed above, such a system would improve the accessibility of these signals to stakeholders, and enable them to more easily select the agents that best match their preferences.

6.3 Theoretical implications

This paper has important implications for several theories relevant to private regulation and corporate social responsibility. With regard to agency theory, it builds on insights from Lin and Starobin and Weinthal that the standard ways for principals to control their agents are often not feasible in corporate social responsibility contexts,[118] and focuses on the more indirect mechanisms that stakeholders use to exercise control over firms. The paper posits that evaluations of credibility and grants of legitimacy are two such mechanisms of control, and shows how these mechanisms connect with each other in a variety of ways. In explaining these mechanisms, the paper contributes to credibility theory by conceptualizing how different forms of credibility – safety, qualification, and dynamism – are embodied in specific signals of credibility – independence, expertise, and transparency – that environmental evaluations send to their principals. It also contributes to legitimacy theory by linking these signals of credibility to different forms of legitimacy, including input and output legitimacy, and pragmatic, moral and cognitive legitimacy.

The paper also has important implications for stakeholder theory, as it posits that different signals of credibility may be particularly salient to specific stakeholder groups, such as civil society organizations, experts, and the general public. These stakeholders consequently may be more willing to grant different forms of legitimacy to programs that send these particular signals. This articulation of the relationship between stakeholder grants of legitimacy and agent signals of credibility contributes to accountability theory by describing a specific example of the dialectic relationship between actors who give account and actors who impose consequences.[119] This framework also builds on the insight from signaling theory

117 Nelson (1970).
118 Starobin and Weinthal (2010); Lin (2012).
119 Painter-Morland (2007); Black (2008).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

that signals can be used to resolve problems of information asymmetry,[120] but raises important questions about the complexity of those signals and the need to make them more accessible to those receiving them, through a voluntary virtual marketplace of standardized signals. The inductive analytic coding process also introduced a more nuanced and multi-dimensional understanding of the concepts of independence, expertise, and transparency. The paper also articulates a hypothesis that given its connections to multiple forms of credibility and legitimacy and salience to the broadest audience, transparency is likely to be the most common characteristic found among these cases.

6.4 Limitations and future research

The data presented in the paper support this hypothesis, but several important caveats about this conclusion should be mentioned. The study's sampling frame was also limited to initiatives existing in 2010 and available in the US, and does not include environmental evaluations only relevant to other countries. The included cases are primarily implemented by evaluation organizations and not firms themselves, which may either reflect their actual proportions in the marketplace or their prominence in online databases, search results, and the academic literature. The data are also counts of coded segments across 245 cases of environmental evaluations, and do not take into account the market share, availability, or relative popularity of individual cases. Thus, they measure the presence or absence of certain traits in these cases, and not the probability that stakeholders will encounter those signals.

The methods used to collect this data also have several limitations. Despite being more rigorous than past efforts to document the landscape of eco-labels and sustainability ratings, the data collected was subject to some coding error. For example, the reliability of the expertise-related data is relatively low, and may be due to the complexity of the concept of expertise or its relatively low prevalence in these cases. While manual coding enables a more nuanced evaluation of text related to such concepts, it is inevitably somewhat subjective.[121]

It is also important to reiterate that while the paper's theoretical framework suggests a set of hypotheses about why these signals of credibility are being sent and how different audiences are responding to them, the data presented in the

120 Nelson (1970).

121 Automated coding was explored as an alternative data collection method, but was found to introduce significantly more errors than a manual approach.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

paper only provides a snapshot of the prevalence of different types of credibility signals, and does not empirically investigate either the causes or effects of these signals. These limitations suggest an important agenda for future research. For example, future research could focus on whether a broader range of firm-led initiatives are sending the same credibility signals as the sample of such cases in this dataset. More work can also be done to explore different stakeholders' perceptions of accountability, credibility, and legitimacy. Interviews with evaluation organization managers could reveal to whom they feel accountable and how they decide what signals of credibility to send. Similar interviews with firm managers could identify how they decide what sustainability claims to make and which signals and stakeholders are most relevant to these decisions.

Surveys could reveal which forms and signals of credibility are most important to consumers and other stakeholders, and the extent to which they grant different forms of legitimacy based on these signals. Experimental designs could determine to what extent different types of credibility result in grants of pragmatic, moral, and cognitive legitimacy from these stakeholders, and what effect the standardization of these signals might have on these perceptions of legitimacy. Given that different groups can perceive credibility depending on the context they are in,[122] it would be useful to investigate these dynamics across different sectors and product categories. Analyzing how perceptions of credibility change over time, respond to specific events, and differ across national contexts would be other valuable lines of research.

It would also be useful to analyze the relationships between signals of credibility – transparency, expertise, and independence – and the organizational connections of environmental evaluation programs. Is independence, for example, more common among programs that are implemented or supported by non-profit organizations? This work could reveal the extent to which signals of credibility are successfully recruiting grants of legitimacy from the stakeholder groups predicted by the conceptual framework presented above. Such grants of legitimacy, especially if they are perceived as enabling more participation by particular stakeholders (e.g., activist groups or large corporations), could have both negative and positive feedback effects on perceptions of these programs' credibility and legitimacy by other groups. Some groups might perceive such increased access as enhancing a program's input legitimacy, while others might perceive it as unfair and strategically problematic for their own interests. Raynolds, Murray and Heller suggest that state involvement, limited public participation, and the exclusion of particular stakeholder groups in some coffee certification programs,

122 Shapin (1995); Fogg and Tseng (1999); Dando and Swift (2003); Alagona (2008).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

for example, fuel concerns about their democratic legitimacy and corporate associations.[123] Future research that takes into account the availability and market share of different programs could reveal the extent to which these dynamics are affecting the popularity of competing programs.

Scholars could also explore alternative explanations for the prevalence of different signals of credibility. The variance in signal levels, for example, may be instead due to differences in difficulty in achieving the associated attributes (e.g., it may be easier to be transparent than independent or qualified). Possessing these attributes therefore may be a signal of overall organizational competence, which may in turn signal greater outcome effectiveness. They may also be proxies for the strength or rigor of the program; following Potoski and Prakash's typology, programs that signal their independence, for example, may be more likely to be "mandarin" programs with stringent standards and strict "long sword" sanctions.[124] Assessing perceptions among both principals and agents of signaling difficulty, outcome effectiveness and program rigor would contribute to testing these alternative theories.

7 Conclusion

Using a rigorously collected dataset on 245 cases of product and corporate environmental evaluations, this paper makes several important theoretical and empirical contributions to our understanding of corporate social responsibility and sustainability assessments. It reveals that, while attempting to solve one principal-agent problem, these assessments present yet another. It explores how expertise, independence and transparency serve as three important signals of credibility for these assessments, and finds that transparency is the most common signal being sent by these programs. The paper posits that such signals of credibility reflect an underlying sense of accountability to particular audiences and an interest in recruiting grants of legitimacy from these audiences. It further suggests that the prevalence of transparency indicate these programs on average are more committed to the public at large than to experts or advocacy organizations. The paper's conceptual framework can assist managers of firms and evaluation organizations that produce environmental evaluations in thinking through how they prioritize different signals of credibility and sources of legitimacy. The empirical data on the prevalence of different signals both improve our understanding of the role

123 Raynolds, Murray, and Heller (2007).
124 Potoski and Prakash (2009: p. 29).

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

of these initiatives as monitors of the environmental and social performance of firms and point toward opportunities for improving their effectiveness as agents of the broader public.

Acknowledgments: The author would like to thank Betsy Albright, Jonas Monast, Shana Starobin, Erika Weinthal and Jonathan Wiener for their comments on a previous version of this paper presented at Duke University's Nicholas School of the Environment Seminar Series.

Funding: Intel People and Practices Research (PaPR).

References

Adams, Carol A. 2004. "The Ethical, Social and Environmental Reporting-Performance Portrayal Gap." *Accounting, Auditing and Accountability Journal* 17 (5): 731–757.

Adams, Carol A., Wan-Ying Hill, and Clare B. Roberts. 1998. "Corporate Social Reporting Practices in Western Europe: Legitimating Corporate Behaviour?" *The British Accounting Review* 30 (1): 1–21.

Alagona, Peter. 2008. "Credibility." *Conservation Biology* 22 (6): 1365–1367.

Auld, Graeme, and Lars H. Gulbrandsen. 2010. "Transparency in Nonstate Certification: Consequences for Accountability and Legitimacy." *Global Environmental Politics* 10 (3): 97–119.

Banerjee, Abhijit, and Barry D. Solomon. 2003. "Eco-Labeling for Energy Efficiency and Sustainability: A Meta-Evaluation of US Programs." *Energy Policy* 31 (2): 109–123.

Bansal, Pratima, and Kendall Roth. 2000. "Why Companies Go Green: A Model of Ecological Responsiveness." *The Academy of Management Journal* 43 (4): 717–736.

Baur, Dorothea, and Guido Palazzo. 2011. "The Moral Legitimacy of NGOs as Partners of Corporations." *Business Ethics Quarterly* 21 (4): 579–604.

Beattie, Vivien, Bill McInnes, and Stella Fearnley. 2004. "A Methodology for Analysing and Evaluating Narratives in Annual Reports: A Comprehensive Descriptive Profile and Metrics for Disclosure Quality Attributes." *Accounting Forum* 28 (3): 205–236.

Beck, A. Cornelia, David Campbell, and Philip J. Shrives. 2010. "Content Analysis in Environmental Reporting Research: Enrichment and Rehearsal of the Method in a British–German Context." *The British Accounting Review* 42 (3): 207–222.

Berlo, David K., James B. Lemert, and Robert J. Mertz. 1969. "Dimensions for Evaluating the Acceptability of Message Sources." *Public Opinion Quarterly* 33 (4): 563–576.

Black, Julia. 2008. "Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes." *Regulation and Governance* 2 (2): 137–164.

Blackman, Allen, and Maria A. Naranjo. 2012. "Does Eco-Certification Have Environmental Benefits? Organic Coffee in Costa Rica." *Ecological Economics* 83: 58–66.

Boiral, Olivier. 2013. "Sustainability Reports as Simulacra? A Counter-Account of A and A+ GRI Reports." *Accounting, Auditing and Accountability Journal* 26 (7): 1036–1071.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

Boiral, Olivier, and Yves Gendron. 2011. "Sustainable Development and Certification Practices: Lessons Learned and Prospects." *Business Strategy and the Environment* 20 (5): 331–347.
Boström, Magnus. 2006. "Establishing Credibility: Practising Standard-Setting Ideals in a Swedish Seafood-Labelling Case." *Journal of Environmental Policy and Planning* 8 (2): 135–158.
Buhr, Nola. 1998. "Environmental Performance, Legislation and Annual Report Disclosure: The Case of Acid Rain and Falconbridge." *Accounting, Auditing and Accountability Journal* 11 (2): 163–190.
Bullock, Graham. 2011. "Green Grades: The Popularity and Perceived Effectiveness of Information-Based Environmental Governance Strategies." PhD Diss., Berkeley, CA: University of California, Berkeley.
Bullock, Graham. 2015. "Independent Labels? The Power behind Environmental Information about Products and Companies." *Political Research Quarterly* 68 (1): 46–62.
Caldwell, Cam, Ranjan Karri, and Pamela Vollmar. 2006. "Principal Theory and Principle Theory: Ethical Governance from the Follower's Perspective." *Journal of Business Ethics* 66 (2/3): 207–223.
Cash, David, William Clark, Frank Alcock, Nancy Dickson, Noelle Eckley, and Jill Jäger. 2002. "Salience, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making." In *John F. Kennedy School of Government, Harvard University, Faculty Research Working Paper Series RWP02-046*. Cambridge, MA: Harvard Kennedy School of Government.
Cashore, Benjamin. 2002. "Legitimacy and the Privatization of Environmental Governance: How Non-State Market Driven (NSMD) Governance Systems Gain Rule-Making Authority." *Governance* 15 (4): 503–529.
Chi, Michelene T. H., Robert Glaser, and Marshall J. Farr, eds. 1988. *The Nature of Expertise*. New York, NY: Psychology Press.
Cho, Charles H., Ronald P. Guidry, Amy M. Hageman, and Dennis M. Patten. 2012. "Do Actions Speak Louder than Words? An Empirical Investigation of Corporate Environmental Reputation." *Accounting, Organizations and Society* 37 (1): 14–25.
Cone Communications. 2012. *2012 Cone Green Gap Trend Tracker*. Accessed August 25, 2013. <http://www.conecomm.com/2012-cone-green-gap-trend-tracker>.
Costa, Sandine, Lisette Ibanez, Maria L. Loureiro, and Stéphan Marette. 2009. "Quality Promotion through Eco-Labeling: Introduction to the Special Issue." *Journal of Agricultural & Food Industrial Organization* 7 (2): 1–6.
"Credibility." 2014. *Merriam-Webster.com*. Merriam-Webster. Accessed April 1, 2010. <http://www.merriam-webster.com/dictionary/credibility>.
Dando, Nicole, and Tracey Swift. 2003. "Transparency and Assurance: Minding the Credibility Gap." *Journal of Business Ethics* 44 (2/3): 195–200.
Darby, Michael R., and Edi Karni. 1973. "Free Competition and the Optimal Amount of Fraud." *The Journal of Law and Economics* 16 (1): 67–88.
Deegan, Craig. 2002. "The Legitimising Effect of Social and Environmental Disclosures – A Theoretical Foundation." *Accounting, Auditing and Accountability Journal* 15 (3): 282–311.
Edelman. 2013. 2013 "Edelman Trust Barometer Finds a Crisis in Leadership." *Edelman*. Accessed January 19, 2015. <http://www.edelman.com/trust-downloads/press-release/>.
Eisenhardt, Kathleen M. 1989. "Agency Theory: An Assessment and Review." *The Academy of Management Review* 14 (1): 57.
Ericsson, K. Anders, and Jacqui Smith, eds. 1991. *Toward a General Theory of Expertise: Prospects and Limits*. Cambridge, UK: Cambridge University Press.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

Farneti, Federica, and James Guthrie. 2009. "Sustainability Reporting by Australian Public Sector Organisations: Why They Report." *Accounting Forum* 33 (2): 89–98.
Fifka, Matthias S., and Maria Drabble. 2012. "Focus and Standardization of Sustainability Reporting – A Comparative Study of the United Kingdom and Finland." *Business Strategy and the Environment* 21 (7): 455–474.
Fogg, B. J., and H. Tseng. 1999. "The Elements of Computer Credibility." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: The CHI Is the Limit*.
Friedman, Milton. 1970. "The Social Responsibility Of Business Is to Increase Its Profits." *The New York Times*, September 13.
Friedman, Monroe. 1999. *Consumer Boycotts: Effecting Change through the Marketplace and the Media*. New York, NY: Routledge.
Frost, Geoffrey R. 2007. "The Introduction of Mandatory Environmental Reporting Guidelines: Australian Evidence." *Abacus* 43 (2): 190–216.
Fung, Archon. 2013. "Infotopia: Unleashing the Democratic Power of Transparency." *Politics and Society* 41 (2): 183–212.
Fung, Archon, Mary Graham, and David Weil. 2007. *Full Disclosure: The Perils and Promise of Transparency*. Cambridge, UK: Cambridge University Press.
Gereffi, Gary, Ronie Garcia-Johnson, and Erika Sasser. 2001. "The NGO-Industrial Complex." *Foreign Policy* (125): 56–65.
Gifford, E. James M. 2010. "Effective Shareholder Engagement: The Factors That Contribute to Shareholder Salience." *Journal of Business Ethics* 92: 79–97.
Gilbert, Dirk Ulrich, and Andreas Rasche. 2008. "Opportunities and Problems of Standardized Ethics Initiatives – A Stakeholder Theory Perspective." *Journal of Business Ethics* 82 (3): 755–773.
Graham, Mary. 2002. *Democracy by Disclosure: The Rise of Technopopulism*. Washington, DC: Governance Institute/Brookings Institution Press.
Gray, Rob, Reza Kouhy, and Simon Lavers. 1995. "Methodological Themes: Constructing a Research Database of Social and Environmental Reporting by UK Companies." *Accounting, Auditing and Accountability Journal* 8 (2): 78.
Hamilton, James T. 2005. *Regulation through Revelation: The Origin, Politics, and Impacts of the Toxics Release Inventory Program*. Cambridge, UK: Cambridge University Press.
Hershey's. Rainforest Alliance Certified. *Hersheys.com*. Accessed November 17, 2014. <http://www.hersheys.com/bliss/our-story/rainforest-alliance.aspx>.
Hess, David. 2007. "Social Reporting and New Governance Regulation: The Prospects of Achieving Corporate Accountability through Transparency." *Business Ethics Quarterly* 17 (3): 453–476.
Hickle, Garth. 2007. "Promoting Product Stewardship with Eco-Labels, Certification Programs, and Product Standards." *Environmental Quality Management* 16 (3): 1–9.
Hovland, Carl I., Irving L. Janis, and Harold H. Kelley. 1953. *Communication and Persuasion*. New Haven, CT: Yale University Press.
Ilchman, Warren Frederick, and Norman Thomas Uphoff. 1969. *The Political Economy of Change*. Berkeley, CA: University of California Press.
Jasanoff, Sheila. 2003. "(No?) Accounting for Expertise." *Science and Public Policy* 30 (3): 157–162.
Jose, Anita, and Shang-Mei Lee. 2007. "Environmental Reporting of Global Corporations: A Content Analysis Based on Website Disclosures." *Journal of Business Ethics* 72 (4): 307–321.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

Kiesel, Kristin, and Sofia B Villas-Boas. 2007. "Got Organic Milk? Consumer Valuations of Milk Labels after the Implementation of the USDA Organic Seal." *Journal of Agricultural and Food Industrial Organization* 5 (1): 1–38.

Koppell, Jonathan G. S. 2005. "Pathologies of Accountability: ICANN and the Challenge of 'Multiple Accountabilities Disorder.'" *Public Administration Review* 65 (1): 94–108.

Krippendorff, Klaus. 2012. *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage Publications, Inc.

Kuklick, Bruce. 1969. "The Analytic-Synthetic and the Descriptive-Evaluative Distinctions." *The Journal of Value Inquiry* 3 (2): 91–99.

Landis, J. Richard, and Gary G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33 (1): 159–174.

Liddick, Don. 2006. *Eco-Terrorism: Radical Environmental and Animal Liberation Movements*. Westport, CT: Greenwood Publishing Group.

Lin, J. 2012. "Governing Biofuels: A Principal-Agent Analysis of the European Union Biofuels Certification Regime and the Clean Development Mechanism." *Journal of Environmental Law* 24 (1): 43–73.

López, M. Victoria, Arminda Garcia, and Lazaro Rodriguez. 2007. "Sustainable Development and Corporate Performance: A Study Based on the Dow Jones Sustainability Index." *Journal of Business Ethics* 75 (3): 285–300.

Ludescher, Jessica Christie, Rubiná Mahsud, and Gregory E. Prussia. 2012. "We Are the Corporation: Dispersive CSR." *Business and Society Review* 117 (1): 55–88.

Mäkinen, Jukka, and Arno Kourula. 2012. "Pluralism in Political Corporate Social Responsibility." *Business Ethics Quarterly* 22 (4): 649–678.

Mattli, Walter, and Tim Büthe. 2005. "Accountability in Accounting? The Politics of Private Rule-Making in the Public Interest." *Governance* 18 (3): 399–429.

Maury, Mary D. 2000. "A Circle of Influence: Are All the Stakeholders Included?" *Journal of Business Ethics* 23 (1): 117–121.

McCubbins, Mathew D., and Thomas Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms." *American Journal of Political Science* 28 (1): 165–179.

McKnight, D. Harrison, Vivek Choudhury, and Charles Kacmar. 2002. "The Impact of Initial Consumer Trust on Intentions to Transact with a Web Site: A Trust Building Model." *The Journal of Strategic Information Systems* 11 (3–4): 297–323.

Mena, Sébastien, and Guido Palazzo. 2012. "Input and Output Legitimacy of Multi-Stakeholder Initiatives." *Business Ethics Quarterly* 22 (3): 527–556.

Mercer, Molly. 2004. "How Do Investors Assess the Credibility of Management Disclosures?" *Accounting Horizons* 18 (3): 185–196.

Milne, Markus J., Sara Walton, and Helen Tregidga. 2009. "Words Not Actions! The Ideological Role of Sustainable Development Reporting." *Accounting, Auditing and Accountability Journal* 22 (8): 1211–1257.

Nadaï, Alain. 1999. "Conditions for the Development of a Product Ecolabel." *European Environment: The Journal of European Environmental Policy (Wiley)* 9 (5): 202–211.

Nelson, Phillip. 1970. "Information and Consumer Behavior." *Journal of Political Economy* 78 (2): 311–329.

Nilsson, Helen, Burcu Tunçer, and Åke Thidell. 2004. "The Use of Eco-Labeling like Initiatives on Food Products to Promote Quality Assurance-Is There Enough Credibility?" *Journal of Cleaner Production* 12 (5): 517–526.

Painter-Morland, Mollie. 2007. "Defining Accountability in a Network Society." *Business Ethics Quarterly* 17 (3): 515–534.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

Parguel, Béatrice, Florence Benoît-Moreau, and Fabrice Larceneux. 2011. "How Sustainability Ratings Might Deter 'Greenwashing': A Closer Look at Ethical Corporate Communication." *Journal of Business Ethics* 102 (1): 15–28.

Patten, Dennis 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.

Paulraj, Antony. 2009. "Environmental Motivations: A Classification Scheme and Its Impact on Environmental Strategies and Practices." *Business Strategy and the Environment (John Wiley and Sons Inc)* 18 (7): 453–468.

Potoski, Matthew, and Aseem Prakash. 2004. "The Regulation Dilemma: Cooperation and Conflict in Environmental Governance." *Public Administration Review* 64 (2): 152–163.

Potoski, Matthew, and Aseem Prakash. 2009. *Voluntary Programs: A Club Theory Perspective*. Cambridge, MA: The MIT Press.

Prakash, Aseem, and Matthew Potoski. 2006. *The Voluntary Environmentalists: Green Clubs, ISO 14001, and Voluntary Regulations*. Cambridge, UK: Cambridge University Press.

Raynolds, Laura T., Douglas Murray, and Andrew Heller. 2007. "Regulating Sustainability in the Coffee Sector: A Comparative Analysis of Third-Party Environmental and Social Certification Initiatives." *Agriculture and Human Values* 24 (2): 147–163.

Raynolds, Laura T., Michael A. Long, and Douglas L. Murray. 2014. "Regulating Corporate Responsibility in the American Market: A Comparative Analysis of Voluntary Certifications." *Competition and Change* 18 (2): 91–110.

Roberts, Robin. 1992. "Determinants of Corporate Social Responsibility Disclosure: An Application of Stakeholder Theory." *Accounting, Organizations and Society* 17 (6): 595–612.

Roe, Brian, and Ian Sheldon. 2007. "Credence Good Labeling: The Efficiency and Distributional Implications of Several Policy Approaches." *American Journal of Agricultural Economics* 89 (4): 1020–1033.

Romzek, Barbara S., and Patricia Wallace Ingraham. 2000. "Cross Pressures of Accountability: Initiative, Command, and Failure in the Ron Brown Plane Crash." *Public Administration Review* 60 (3): 240–253.

Rondinelli, Dennis, and Gyula Vastag. 2000. "Panacea, Common Sense, or Just a Label?: The Value of ISO 14001 Environmental Management Systems." *European Management Journal* 18 (5): 499–510.

Rowley, Tim, and Shawn Berman. 2000. "A Brand New Brand of Corporate Social Performance." *Business and Society* 39 (4): 397–418.

Sadowski, Michael, Kyle Whitaker, and Frances Buckingham. 2010. "Rate the Raters Phase Two: Taking Inventory of the Ratings Universe." SustainAbility. Accessed 5 December 2010. <www.sustainability.com/library/rate-the-raters-phase-two>.

Scharpf, Fritz W. 1997. "Economic Integration, Democracy and the Welfare State." *Journal of European Public Policy* 4 (1): 18–36.

Shapin, Steven. 1995. "Cordelia's Love: Credibility and the Social Studies of Science." *Perspectives on Science* 3 (3): 255–275.

Sim, Julius, and Chris C. Wright. 2005. "The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements." *Physical Therapy* 85 (3): 257–268.

Smith, N. Craig, Guido Palazzo, and C. B. Bhattacharya. 2010. "Marketing's Consequences: Stakeholder Marketing and Supply Chain Corporate Social Responsibility Issues." *Business Ethics Quarterly* 20 (4): 617–641.

Spence, Michael. 1973. "Job Market Signaling." *The Quarterly Journal of Economics* 87 (3): 355–374.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library

Starobin, Shana, and Erika Weinthal. 2010. "The Search for Credible Information in Social and Environmental Global Governance: The Kosher Label." *Business and Politics* 12 (3): 1–35.
Suchman, Mark C. 1995. "Managing Legitimacy: Strategic and Institutional Approaches." *Academy of Management Review* 20 (3): 571–610.
Sustainable Commodity Initiative. 2010. "The State of Sustainability Initiatives: A Window into the World of Voluntary Sustainability Initiatives." Accessed December 2, 2010. <http://sustainablecommodities.org/ssi/update1>.
Turban, Daniel B., and Daniel M. Cable. 2003. "Firm Reputation and Applicant Pool Characteristics." *Journal of Organizational Behavior* 24 (6): 733–751.
U.S. Federal Trade Commission. 2010. *Proposed Revisions to the Green Guides*. Available from: <http://www.gpo.gov/fdsys/pkg/FR-2010-10-15/html/2010-25000.htm> Accessed December 5, 2010.
U.S. Green Building Council. *Distinguishing Your Expertise*. Accessed April 27, 2014. <http://www.usgbc.org/leed/credentials>.
UL Environment and Shelton Group. 2014. "Claiming Green: The Influence of Green Product Claims on Purchase Intent and Brand Perception." Available from: <http://environment.ul.com/claiminggreen/> Accessed December 20, 2014.
Uphoff, Norman. 1989. "Distinguishing Power, Authority & Legitimacy: Taking Max Weber at His Word by Using Resources-Exchange Analysis." *Polity* 22 (2): 295–322.
Vogel, David. 2005. *The Market for Virtue: The Potentials and Limits of Corporate Social Responsibility*. Washington, DC: Brookings Institution Press.
Williams, Cynthia A. 1999. "The Securities and Exchange Commission and Corporate Social Transparency." *Harvard Law Review* 112 (6): 1197–1311.
Wood, Dorothy, and Donald G. Ross. 2006. "Environmental Social Controls and Capital Investments: Australian Evidence." *Accounting and Finance* 46 (4): 677–695.
World Resources Institute and Big Room. 2010. "Global Ecolabel Monitor." Accessed December 1, 2010. <http://www.wri.org/publication/global-ecolabel-monitor>.

Supplemental Material: The online version of this article (DOI: 10.1515/bap-2014-0028) offers supplementary material, available to authorized users.

- 10.1515/bap-2014-0028
Downloaded from PubFactory at 08/15/2016 08:25:16PM
via Davidson College Library