International Journal of Accounting Information Systems 23 (2016) 28–44

Contents lists available at ScienceDirect

International Journal of Accounting Information Systems

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

How measurement framing and accounting information system evaluation mode influence environmental performance judgments☆

Hank C. Alewine *, Christopher D. Allport, Wei-Cheng Milton Shen

College of Business Administration, University of Alabama in Huntsville, Huntsville, AL 35899

ARTICLE INFO

Article history:
Received 5 October 2015
Received in revised form 5 August 2016
Accepted 6 October 2016
Available online 24 October 2016

Keywords:
Environmental accounting information systems
Measurement framing
Decision making
General evaluability theory
Evaluation mode

ABSTRACT

How does information presentation within an accounting information system (AIS) influence environmental performance judgments? Decision makers generally analyze alternatives' performances in one of two evaluation modes: jointly or separately. Joint mode provides greater measure evaluability because of available comparisons between alternates. Thus, additional information garners greater decision weight in separate mode, where less contextual information exists. However, many environmental decision settings use separate evaluation mode because of no viable alternatives (e.g., large pollution abatement investments). In this setting, General Evaluability Theory (GET; Hsee and Zhang, 2010) suggests low measurement evaluability when low measurement knowledge and non-inherently understood measures exist—both common characteristics in environmental settings. This study introduces attribute framing to the GET framework as important to consider when analyzing environmental decision differences across modes, because frames are often a necessary component of information presentation and different descriptions often lead to different decisions (Dunegan, 1993). Experimental participants ($n = 206$) evaluated factory environmental performances with joint/separate mode and positive/negative attribute framing. Findings inform AIS designers as results suggest evaluation mode moderates the presentation of attribute frames on decisions. Specifically, higher (lower) evaluations occur when using positive (negative) framing, and this effect is more (less) pronounced in separate (joint) mode. Findings also suggest that more consistent judgments occur across evaluation mode with positive compared to negative framing of performance measures.

1. Introduction

Accounting information systems (AIS) are designed to collect and store data and to present this data to decision makers as relevant and reliable information (Romney and Steinbart, 2015). Decision makers evaluate this information in a wide range of organizational settings. Increasingly, AIS data matrices are expected to include nontraditional, nonfinancial, and relatively unfamiliar measurements. Decision settings consisting of large proportions of this type of accounting information include fair value accounting (Benston, 2006), goodwill accounting (Wines et al., 2007), and performance evaluations via a balanced scorecard (Humphreys

☆ We thank the University of Alabama in Huntsville for financial support, Tim Miller for assistance with data collection, and anonymous reviewers for insightful feedback.

* Corresponding author.
E-mail addresses: hank.alewine@uah.edu (H.C. Alewine), allporc@uah.edu (C.D. Allport), milton.shen@uah.edu (W.-C.M. Shen).

http://dx.doi.org/10.1016/j.accinf.2016.10.002
1467-0895/

and Trotman, 2011; Kaplan and Wisner, 2009). This decision setting is particularly prominent with environmental accounting information, which entities increasingly consider in decisions aimed at improving environmental stewardship strategic objectives (Joshi et al., 2001; Simnett et al., 2009). Because of the increasing focus afforded to the collecting, processing, and presenting of environmental accounting information for communication to entity stakeholders and for assessment to aid in environmentally conscious decisions, this study focuses on decision settings involving environmental AIS (Brown et al., 2005; Dilla and Steinbart, 2005).

When dealing with nontraditional or unfamiliar measurements, a particular challenge for an environmental AIS is to *present* environmental information in such a way that it alleviates, rather than accentuates, the cognitive hurdles decision makers will inevitably face when evaluating this data (Alewine, 2010; cf. Stone and Schkade, 1994). When dealing with the evaluability of any attribute such as a performance measurement, Hsee and Zhang's (2010) General Evaluability Theory (GET) suggests that three characteristics should be considered – the nature of the attribute, the knowledge about the attribute, and the mode of attribute presentation. An attribute's "evaluability" refers to the degree of ease in which a decision maker can successfully assess the information that is attached to the attribute, resulting in effective decision making. Low evaluability results when all three attribute characteristics are low (not instinctively or naturally evaluable, limited knowledge, and an isolated evaluation mode). The nature and knowledge characteristics are often low for attributes in environmental accounting settings, and thus the presentation mode garners extra significance in environmental AIS design when considering how to limit possible decision biases. Joint evaluation mode refers to a simultaneous analysis of multiple alternatives under consideration in a decision. Separate evaluation mode refers to the analysis of one alternative in isolation for a decision.

Based on the GET framework, evaluation mode is expected to play an important role in minimizing decision bias in an environmental AIS; however, other presentation factors are likely to impact attribute evaluations. One likely example would be attribute framing. Attribute framing occurs when functionally equivalent information is described either positively (chance of success) or negatively (chance of failure). The attribute framing effect occurs when the positive description of an item is evaluated more highly than the same item described with a negative description (Levin et al., 1998). The attribute framing effect has been robust across various decision contexts (medical vs. business) and decision makers (expert vs. novice), and the characteristics of environmental accounting information are likely to require the use of positive and/or negative descriptive valences (see Levin et al., 1998 for a literature review).

While the environmental context likely requires attribute frames, any AIS is likely to include characteristics to limit any decision bias based on positive or negative attribute descriptions. Specifically, attribute descriptions in an AIS are unlikely to be as valence-charged as those seen in psychological research; examples such as life vs. death are certainly not likely, while 'weaker' descriptors such as recycled vs. *not* recycled are likely. Such descriptive differences have been shown to impact the size of framing effects (Kuhberger, 1995). In addition, an AIS would likely include a benchmark as a reference point for decision makers. Very few framing studies have considered benchmarks (see Kerler et al., 2014 for an exception). Most importantly for the current study, however, is the impact of presentation mode on attribute framing effects. We predict that joint evaluations will increase attribute evaluability in decision making and therefore limit the effect of positive and negative attribute frames. Overall, the current study seeks to analyze the impact of *realistic* attribute frames on environmental accounting decisions across separate and joint evaluation modes.

In a laboratory experiment, participants ($n = 206$) were randomly assigned to one of four between-participant conditions; manipulated variables include attribute framing (positive, negative) and evaluation mode (separate, joint evaluation). An industrial benchmark was included as a control variable, with assessed factories being either all better-performing or all worse-performing than the benchmark (a within-participant variable). Participants assessed two factory performances on two environmental measures and evaluated them based on their environmental performance. This process was repeated for three other pairs of factory evaluations.

As predicted, the attribute framing effect was robust in expected directions in the separate evaluation mode—higher (lower) evaluations occurred when using positive (negative) attribute frames. These differences persisted even in the presence of moderate attribute descriptions and the presence of attribute benchmarks. However, consistent with the literature on presentation mode in decisions, when providing joint evaluations along with the industry benchmark, the attribute framing effect was not observed. That is, no evaluation differences were found between settings using positive and negative attribute frames.

Various information presentation characteristics have been examined by information system or AIS researchers (Kelton et al., 2010). The study's reported results inform decision makers and environmental AIS designers on how presentation mode (separate vs. joint evaluation) can benefit decision makers faced with non-traditional and/or unfamiliar data. Specifically, this study provides insight on the advantage of joint evaluation comparisons over separate evaluation (when possible), as descriptive valence differences do not impact decisions in joint evaluations. Additional analysis also suggests more consistent judgments occur across evaluation mode when positive descriptions are used compared to negative descriptions. This finding suggests additional presentation implications, as positive descriptions may cause less cognitive decision bias than negative descriptions. Finally, this study answers calls in the AIS literature to proactively address research matters that combine AIS and management accounting fields, and to expand perspectives on dealing with such issues by introducing theoretical foundations, such as GET, that have not been traditionally considered in the AIS literature (Granlund, 2011; Vasarhelyi, 2012; see Rom and Rohde, 2007).

This paper next describes the hypotheses development, followed by the procedure and reported results of an experiment. The paper concludes with a discussion of contributions, limitations, and avenues for future research inquiries.

1.1. Literature and hypotheses

1.1.1. Relevant AIS and environmental accounting literature

To achieve organizational goals, accounting information systems (AIS) collect, process, and store data to produce relevant, reliable, complete, and timely information. This information is then *presented* to decision makers in a useful and intelligible format (Romney and Steinbart, 2015). Empirical research has evidenced that information presentation has a significant impact on decision making. Some research has focused on the relative advantages of a presentational format (e.g., graphs vs. tables; paper vs. computer screen; 2D vs. 3D), others on the influences of a presentational element (e.g., colors, graph slopes, gridlines, interactivity, multimedia), still others on the interaction between information presentation and other factors, such as user characteristics or task requirements (Bacic and Appan, 2012; Debreceny, 2011; Dilla et al., 2010; Kelton et al., 2010; Shaft and Vessey, 2006; Vessey, 1991).

The role of information presentation in an AIS is particularly relevant considering the changing data requirements and expectations of organizational decision makers and stakeholders. The modern AIS is expected to collect not only traditional financial data, but also non-traditional and/or nonfinancial data for any number of initiatives, including enterprise resource planning, supply chain management, customer relationship management, or balanced scorecard assessment (Dilla et al., 2010; Gray and Bebbington, 2001). In addition to these initiatives, the current study focuses on another increasingly important source of organizational data, environmental accounting data. Considering the challenges inherent in utilizing non-traditional data regarding the environmental impact of organizations, how to effectively present this data becomes an important issue for environmental AIS designers and users.

The green information system literature considers two broad areas of environmental impact (see Jenkin, Webster, and McShane 2011). One approach is to analyze how information systems directly impact the environment, such as direct energy savings stemming from implementing new environmentally conscience information technology (Cater-Steel and Tan, 2010). This study focuses on the second area, how information systems can indirectly impact the environment, such as with the implementation of an environmental AIS to efficiently provide data that helps to improve environmental management decisions. The literature notes that many AIS do not provide sufficient environmental information for decisions (Jasch and Savage, 2008), and only sparse tools exist to manage such information (Weigand and Elsas, 2012). To complicate matters, field evidence also suggests that when environmental measurement systems indeed exist, they tend to operate under the purview of environmental management employees without adequate accounting expertise input (Albelda, 2011); information system units are often not part of entities' environmental analyses (Huang, 2009).

The above discussed research may help to explain why the literature documents operational and cultural challenges for successfully developing and integrating environmental AIS for more effective environmental decisions (see Blackburn et al., 2014; Dillard, 2008). Political, financial, and stakeholder pressures heavily influenced environmental objectives relating to a local government's adoption of environmental measures in its policy making process (Ball, 2005). Similar factors derailed an introduction of a social and environmental information system in a fair trade organization (Dey, 2007), as well as an attempted implementation of environmental valuation methods in a governmental forest management agency (Herbohn, 2005). Overall, though, the literature has not thoroughly explored exactly how entities may provide more inclusion of information system perspectives in environmental management concerns (Jenkin, McShane, and Webster 2011). Thus, many opportunities currently exist for both practitioners and researchers to better properly develop information systems that meet environmental management needs. To better accomplish this objective, we analyze characteristics of environmental accounting information.

Environmental accounting information usually involves unfamiliar, nonfinancial, and nontraditional characteristics (see Schaltegger and Burritt, 2010). The literature recognizes that such information increasingly impacts judgments, and particularly environmental judgments (see Clarkson et al., 2004; Dikolli et al., 2009; Hughes, 2000; Joshi et al., 2001; Smith, 2002; Sriramachari and Chandra, 1997). Indeed, the nature of these data results in many challenges for their capture, measuring, and reporting for information users (Brown et al., 2005). This occurs especially since the conversion of nonfinancial environmental measures to financial measures is often very noisy or undesirable (Epstein, 2008; Gray and Bebbington, 2001), while more complicated “unique” measures are more cognitively expensive to process effectively (Stone and Schkade, 1994). The above mentioned challenges underscore the need to devise better methods for compiling and propagating environmental information in order to better achieve environmental stewardship strategic objectives. Effective environmental AIS must be developed to address these challenges so that entities can better capture, measure, and report environmental accounting information for a variety of initiatives, such as communications to entity stakeholders, as well as assessments to aid in environmentally conscious decisions (Brown et al., 2005; Dilla and Steinbart, 2005; see Marelli, 2015).

1.1.2. General evaluability theory

Features of environmental accounting information present challenges when considering decision consequences of cognitively processing this contemporary information with respect to data presentation. Thus, this study considers the challenges of presenting environmental accounting information in AIS using Hsee and Zhang's (2010) General Evaluability Theory (GET; see Fig. 1 for an overview), especially since this theory provides guidance on analyzing presentation impacts on performance measurements' evaluability for measures containing characteristics often found in environmental accounting information. This cognitive psychological framework suggests that attributes (e.g., environmental performance measures) vary in evaluability based on three factors—mode, knowledge, and nature. Evaluation mode includes evaluating alternatives (e.g., managers, business units, or investment opportunities) either together or in separate independent cases, known as joint and separate mode, respectively. Knowledge

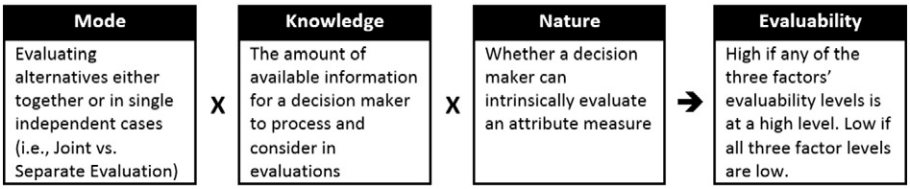

Fig. 1. General evaluability theory framework overview.

refers to the amount of information a decision maker brings to the attribute analysis (i.e., their level of expertise on the subject matter). Finally, nature refers to whether a decision maker can intrinsically evaluate an attribute measure.

Research involving various components of the GET framework (mode, knowledge, and nature factors) has been employed in multiple disciplines, such as economics (e.g., Ariely et al., 2003), social and cognitive psychology (e.g., Hsee and Zhang, 2004; Morewedge et al., 2009), marketing (e.g., Hsee et al., 2009), and recently in accounting (Alewine and Stone, 2016). Some research focuses on the three factors' impact on measurement evaluability—evaluation mode (Ariely et al., 2003; Hsee and Zhang, 2004; Kogut and Ritov, 2005), measurement knowledge (Ariely and Loewenstein, 2000; Morewedge et al., 2009; Yeung and Soman, 2007), or on the "nature" of information (Hsee et al., 2009). Also, research has explored interactions between two of the three factors in impacting the evaluability of data, including analysis of mode with either knowledge (Hsee et al., 2013) or nature (Hsee et al., 2009). This study will explore changes to the mode factor while holding the other two factors constant, as discussed in more detail later in this section.

Attributes can be assessed at either a high or low evaluability level for each of the three factors. According to the framework, only one of the three factors has to be assessed at a low level for the information to be deemed to have low evaluability. Examples of low and high levels of evaluability for each factor follow. For the mode factor, evaluating multiple alternatives for an environmental investment decision simultaneously (i.e., joint evaluation) garners high evaluability, while investment evaluations in isolation or asymmetrically (i.e., separate evaluation) provide low evaluability levels. This occurs because attributes evaluated in a joint evaluation mode contain more context in which to evaluate the information; at a minimum, one may compare alternatives' attribute performance values. Such a comparison is not available in separate evaluation mode, and less context there leads to relatively less evaluability for the attributes. As an example for illustrating the two evaluability levels for the knowledge factor, consider that chief financial officers evaluate a return on investment value with high knowledge levels, thus making the measure highly evaluable. However, decision makers evaluating environmental performance measures often contain rather limited knowledge (see KPMG, 2011), leading to a low knowledge evaluability level for those environmental measures. To illustrate high evaluability for the nature factor, consider happiness—people inherently understand whether they are happy or not happy. On the other hand, the measure "biodiversity index" is not intrinsically understood and therefor contains a low evaluability level for the nature factor.

In seeking to understand the challenges presented by environmental accounting information, the GET framework provides valuable insight. Specifically, considering the attributes of environmental accounting information as having a low evaluability provides AIS designers a better understanding of challenges and remedies related to the capture and reporting of meaningful environmental information. This is especially valuable, as despite best efforts from very bright accountants and managers, assessing environmental measures likely involves low knowledge pertaining to the measures, and the measures are not intrinsically evaluable (e.g., a low assessed nature factor). In fact, practitioners indicate they make more decisions with financial performance measures compared to nonfinancial measures because financial measures are more familiar, common, and understood (Abdallah and Alnamri, 2015). Common performance measures tend to be weighted more in decisions compared to more unique measures (Lipe and Salterio, 2000). Not only do many environmental measures contain both nonfinancial and unique characteristics that result in less management familiarity for the measures compared to traditional financial measures, but also a KPMG (2011) report discusses challenges relating to deficiencies in environmental skillsets and experiences for business practitioners using environmental accounting information. Thus, environmental performance measures likely contain low levels of evaluability with respect to both the knowledge and nature factors. Ultimately, this suggests that evaluation mode determines the environmental measures' information evaluability—joint mode leads to high evaluability, while a separate mode evaluation results in low information evaluability.

Alewine and Stone (2016) introduce the accounting literature to the GET framework while exploring how environmental performance assessments are influenced by evaluation mode (either joint or separate) and the signal given by comparing performances with benchmark information (either positive or negative). By using environmental performance measures for the evaluations, low levels of the knowledge and nature factor existed (discussed above), so they were able to explore how changes in evaluation mode [using joint (separate) mode for a high (low) evaluability setting] interacted with changes in the signal given by comparing the alternatives' environmental performances to benchmark information [a positive (negative) signal when performances were better (worse) than the benchmark]. They found that performance decisions for two options under review (factory environmental performances) differed across evaluation mode when benchmark information indicated that both options performed worse than the benchmark—lower evaluations were given in separate mode. However, decision differences across evaluation mode did not occur when the benchmark information indicated that both options instead performed better than the benchmark. Thus, Alewine and Stone (2016) extend the environmental accounting literature by providing insight on how

mode and benchmark signaling interact to influence environmental performance evaluations.[1] They also extend GET to consider benchmark signal valences when analyzing changes in evaluation mode.

1.1.3. Attribute framing

While GET highlights the importance of presenting data for effective evaluation, the framework is silent with regard to another presentation consideration that could influence environmental decision making, namely attribute framing. Attribute framing occurs when a performance measure can be presented either positively (success) or negatively (failure) while maintaining its normative value (Levin et al., 1998). Research across multiple domains has shown differences in performance measure evaluations when positive and negative attribute frames are compared, with positive frames receiving higher evaluations than negative frames (Kerler et al., 2012; Brockner et al., 1995; Zhang and Buda, 1999). Levin and Gaeth (1988) provide a commonly cited example of attribute framing by describing hamburger meat as either 75% 'lean' or 25% 'fat,' and they find that quality ratings are higher for the 'lean' meat. Levin et al. (1998) propose that these effects occur due to associations made with like experiences, either positive or negative, that are generalized in the evaluation. The same description valence decisions discussed above are necessary for environmental performance measures, as environmental goals may be either partly *satisfied* or *not satisfied*, a new heating system could either partly *save* or *lose* energy, a percentage of scrap is either *recycled* or *not recycled*, and energy observers could either *compliment* or *criticize* aspects of performances.

The signal garnered from the way in which performance attributes are framed will be present in environmental performance judgments. In fact, the threat from possible divergent attribute frames may be of particular concern when dealing with judgments involving environmental accounting information for two reasons. First, nonfinancial environmental information reporting is currently likely to be less standardized when compared to traditional financial accounting reports, providing greater variation in item descriptions/measures. In fact, qualitative studies and surveys suggest that nonfinancial information may be considered more valuable than financial information in an environmental accounting context (Albelda, 2011; KPMG, 2011; KPMG, 2015), with some researchers questioning any attempts to convert nonfinancial information to financial due to the potential loss of information richness and insight (Gray and Bebbington, 2001). Second, because environmental accounting information often has low evaluability characteristics, any information or information bias provided in environmental judgments will likely carry more decision weight in the attribute evaluation. In such cases, the valence association provided by the attribute frame is more likely to be anchored on, resulting in a greater chance of decision bias. Thus, considering framing effects on environmental accounting information processing will better inform environmental AIS designers and other users of this information.

Because attribute framing effects have not been considered in a GET framework and the non-traditional nature of environmental information, we formally present a hypothesis that serves as a reasonableness check regarding the impact of positively and negatively framed attributes.

H1. Higher (lower) evaluation judgments will occur when assessing environmental performance measures that contain positive (negative) framing.

1.1.4. Attribute framing and evaluation mode

Based on the above hypothesis, we expect attribute framing to impact decisions with environmental accounting information. It is at this point that GET provides particularly helpful insight for environmental AIS settings. As previously stated, environmental accounting information often has low evaluability, but GET provides a finer delineation of evaluability into knowledge, nature, and mode. So in an environmental AIS setting, a decision maker may have low levels of knowledge for environmental accounting information and the environmental measures may not be naturally or inherently evaluable (both common characteristics of environmental accounting decision settings), but the AIS designer can still present either the joint or separate evaluation mode. GET suggests that even though the environmental accounting information is the same across evaluation mode, decisions using the information may differ because in joint mode, the attributes will be more evaluable than in separate mode (see O'Donnell and David, 2000). The difference between the two evaluation modes occurs because the evaluation occurring in joint mode contains more context—there is another alternative that can be used as a comparative point. This context is not available in separate mode (see Mălăescu and Sutton, 2015 for insight on how restrictive systems influence decisions). As a result, the evaluation in joint mode is more evaluable than in separate mode, leading to different decision sensitivities across evaluation mode. Hence, joint mode increases the contextuality of the unfamiliar information and, therefore, provides the decision maker with more opportunity to understand the substance of the information over the form.

The above discussion suggests that increasing the evaluability of information through a joint evaluation mode could serve to mitigate any attribute framing effect. One may wonder why any decision would be considered in separate evaluation mode given the information limitations in assessing information in that mode without adequate supplemental evaluability information. However, environmental decisions may exist in settings where only one viable alternative currently exists. An example of this setting is a unique capital intensive decision on investing in scrubber technology because factors such as a factory's location or substantive technological conditions with the factory render other options unviable (Kaminski, 2003). In this realistic setting, an entity will have to decide if the one alternative warrants near-term investment, or whether the entity is willing to wait for a certain

[1] See Bonner (2008) for an extensive review of presentation format studies. However, Bonner notes, "It is impossible to summarize the findings of presentation format studies in any meaningful way" (p. 193) because of the wide variety of research topics and approaches in the literature. However, to the best of our knowledge, Alewine and Stone (2016) seem to be the only accounting study to address evaluation mode issues in the context of the current study.

time period for technological advances to occur that would allow an assessment to become more favorable for consideration later, either by the original option improving itself or by technological advances and market opportunities adding more alternatives as viable options for the decision (and thus shifting the evaluation mode to joint). Thus, analysis of environmental information in joint evaluation mode may not always be possible in decision settings, resulting in a need to better understand settings in which decisions differ between the two modes.

From the perspective of the attribute framing literature, joint evaluations provide additional reference points for the decision maker to consider in interpreting information. Attribute frames have been explained as simple valence-based biases, not dependent on reference points, but Schweitzer (1995) provides an example of multiple reference points limiting the effect of attribute framing. In addition, Sullivan and Kida (1995) find that multiple reference points can play a key role in corporate investment decisions and that multiple reference points, particularly conflicting reference points, should be further investigated. We expect that information from alternatives in joint evaluations, representing additional reference points, will be given more decision weight than a valence-based attribute frame, and therefore, joint evaluation modes will moderate the effect of attribute framings.

H2. Evaluation mode will moderate environmental performance measurements' framing valence's effect on judgments. Specifically, the difference in judgments from using positive and negative measurement framing will be greater in separate than in joint evaluation mode.

2. Method

To test the above hypotheses, participants were randomly assigned to one of four between-participant conditions; manipulated variables included evaluation mode (joint, separate) and environmental measurement framing (positive, negative). Participants assessed four pairs of factories using two environmental measures and judged how well the factories achieved environmental stewardship objectives. Information on the participants and experimental procedure follow.

2.1. Participants

Student participants ($n = 206$) were recruited from a large business college at a public university in the United States. The experiment was conducted using paper and pencil. As an incentive for participation, participants were offered both extra class credit and chances to win in drawings for $25 gift cards.[2] Student participants are appropriate for this study because of the cognitive psychological reasoning for the study's addressing of the research question. This decision is externally valid because the literature demonstrates that practitioners shy away from nonfinancial measures in performance measurement systems due to lack of familiarity (Abdallah and Alnamri, 2015), and practitioners using environmental accounting information for decisions often lack environmental expertise (KPMG, 2011); thus, in this study's context, students and practitioners likely share similar judgment and decision making processes. Also, this research decision allows for the use of more experienced professional participants, a valuable research resource, in research projects that truly require their insight to successfully address certain research questions (Libby et al., 2002).

2.2. Procedure

Fig. 2 shows the study's experimental procedure. Participants assumed the role of a hypothetical manager tasked with assessing factories based on their environmental performances. For a pair of factories, two environmental performance measures were provided, along with industrial benchmark values (mean and range) for each of the measures. Depending on which evaluation mode participants were randomly assigned to, participants either analyzed a pair of factories together (joint mode) or one at a time (separate mode). This process was repeated for three other pairs of factories. To avoid demand effects, participants assigned to separate evaluation mode assessed one factory from each of the four factory pairings before assessing the second factory from each pairing.

Participants were informed that the two performance measures were equally important. Participants were told to indicate how much they would be willing to pay to invest in each factory on a scale of one million to a hundred million dollars; the better the factory performed with respect to the environmental measures, the higher their willingness to pay should be. Factories performed similarly; each factory in the pairing was favored by one of the two measures. Two of the pairings contained factory performances that were better than the performance measures' benchmark values, and the other two factory pairs contained performances that were worse than the benchmark values. See Table 1 for the environmental performance measures and their values for each of the four factory pairings.

After assessing a factory, participants were asked Likert-scaled questions pertaining to their confidence in their decision, as well as how well they understood each of the measures. This entire process repeated for all four factory pairings, resulting in eight total judgments. See the Appendix for a sample research instrument. To mitigate potential order effects driving the results,

[2] Permission was sought and granted from the university's Institutional Review Board to conduct research using human participants in an experimental setting.

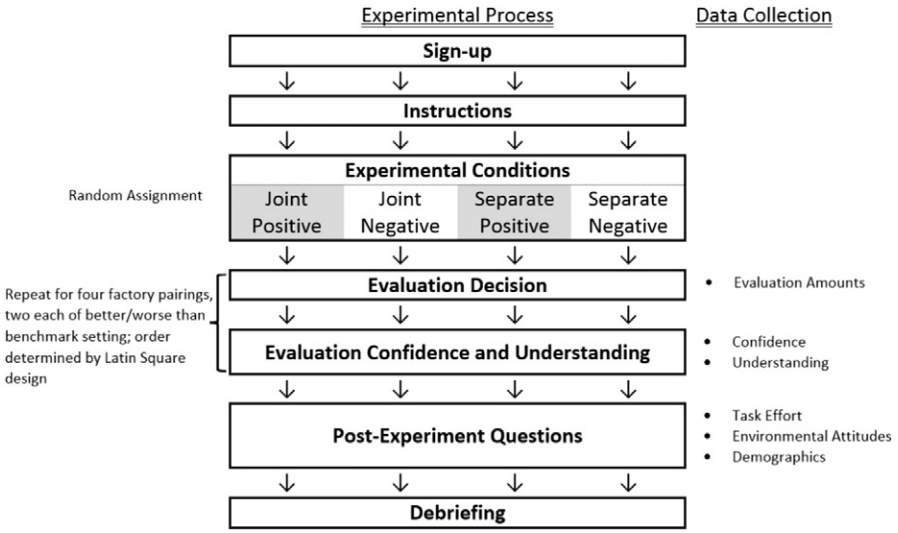

Fig. 2. Experimental process.

participants assessed factories in varying order. Table 2 shows the Latin Square design that illustrates the four sequences of factory pairings used in this study. After participants assessed all factories, they answered post-experimental Likert-scaled questions on task effort and difficulty, as well as on environmental attitudes. Finally, participants answered demographic questions and signed up to enter multiple drawings for $25 gift cards.

Table 1
Summary of instrument's environmental performance measures.

Pairing ID	Metric	Frame	Range Low Value (50% from mean)	Worse-than-reference data points		EI (mean value)	Better-than-reference data points		Range High Value (50% from mean)
				40% from Mean	20% from Mean		20% from Mean	40% from Mean	
1	Percentage of scrap materials [recycled/not recycled] back into production	Positive	30	35	47	59	71	83	89
		Negative	71	65	53	41	29	17	12
	Of the total compliments and complaints given by customers on environmental performance, the percentage of [compliments/complaints]	Positive	17	20	27	34	41	48	51
		Negative	83	80	73	66	59	52	49
2	Percentage of energy [saved/lost] when using a new heating system	Positive	32	38	51	64	77	90	96
		Negative	68	62	49	36	23	10	4
	Percentage of future strict water regulatory standards currently [satisfied/not satisfied]	Positive	22	26	35	44	53	62	66
		Negative	78	74	65	56	47	38	34
3	Of the total positive and negative media stories on the company's environmental performance, the proportion of [positive/negative] media stories	Positive	20	23	31	39	47	55	59
		Negative	81	77	69	61	53	45	42
	Percentage of production byproducts that [can/cannot] be reused for other purposes	Positive	32	38	51	64	77	90	96
		Negative	68	62	49	36	23	10	4
4	Percentage of sulfur dioxide emissions that are [filtered from release/released] into the air when using a new air pollution control device	Positive	27	32	42	53	64	74	80
		Negative	74	68	58	47	36	26	21
	Percentage of future stringent environmental quality standards currently [achieved/not achieved]	Positive	23	28	37	46	55	64	69
		Negative	77	72	63	54	45	36	31

Note: The supplemental evaluability information provided is the mean of the measure's high and low range values. The values are either 20% or 40% above and below the mean value, and range values are about 50% from the mean. Many of the mean values reflect practitioner statistics from various sources. The better-than-reference within-participant variable consists of Pairings 1 & 3, while Pairings 2 & 4 give worse-than-reference information. For counterbalancing purposes, 1) in each pairing, one measure's mean value is higher with the positive frame, and the other measure's mean value is higher with the negative frame, and 2) Factory A performs better on the first measure in Pairings 1 & 2, and vice versa for Pairings 3 & 4. The sum of positive and negative-framed percentages may not add exactly to 100% for each measure because of rounding.

Table 2
Random balance experimental design.

Trial	Pairing
1	1/2/3/4
2	2/3/4/1
3	3/4/1/2
4	4/1/2/3

Note: Participants were randomly assigned to 1 of 4 trials. Pairing number refers to the Pairing ID used in Table 1. The pairing order was determined by a Latin Square design. There was no reason to expect potential order effects from the presentation of only two factories and their two measures. Thus, the presentation order of factories and their measures was held constant. In separate evaluation mode conditions, participants viewed one factory for each pairing at a time. They evaluated the first factory from all pairing s before evaluating the second factory from each pairing. The second factories were presented in the same pairing order as the first factories.

2.3. Necessity of benchmark information as a within-participant variable

Each environmental performance measure contained benchmark information in the form of industrial means and ranges. Benchmark information was included to allow for basic context in separate evaluation mode.[3] Benchmark information is necessary because without this added context, participants in separate mode would evaluate environmental performances with no context, rendering their responses too noisy and unrealistic for reasonable practitioner settings, despite the necessity for experimental designs to contain artificial components that sacrifice external validity (see Shadish et al., 2002).

Thus, providing benchmark information in the study allows for enhanced external validity, but careful consideration had to be given to the type of benchmark information included in the instrument so that the experimental design could still rigorously investigate the research question without confounds. Specifically, consider the valence that is given when one compares two alternatives' performance values with a benchmark; there are three possible outcomes (Hsee and Leclerc, 1998). First, the benchmark performance value may fall between the two alternatives' performances. In this scenario, a participant making a judgment in separate mode would give a better judgment for the alternative performing better than the benchmark, and a worse judgment for the alternative performing worse than the benchmark. These judgments would be qualitatively similar if they were made in joint mode, despite the relatively extra context available in joint mode (i.e., comparison between the alternatives). Since the current study focuses on settings in which decisions may differ across evaluation mode, the experimental design does not provide benchmark information that yields this particular outcome.

Instead, this study focuses on the other two possible outcomes. Specifically, both alternatives may perform better than the benchmark performance, or both alternatives may perform worse than the benchmark performance. This design decision was selected so that the benchmark information would provide external validity (especially in separate mode), but not so much guidance that differences in judgment would not be detected across evaluation mode (such as in the first outcome discussed above). From a framing perspective, a benchmark could limit the decision weight appropriately given to the attribute frame. However, past findings on attribute framing effects still suggest that decisions will be biased in predictable ways. Statistical tests reported later will control for the benchmark information when analyzing evaluation mode effects as a moderating variable for framing effects on judgments of environmental performance measures.

2.4. Measurements

As indicated previously, research participants evaluate a total of eight factories' environmental performances based on information presented in different evaluation modes and with different performance measurement framing valences. Thus, the study's outcome variable is the willingness to invest amount as an indicator of the factory's performance, whereas manipulated variables are evaluation mode and performance measurement framing valence.[4]

To ensure comparability across experimental conditions and to mitigate potential confounds of effort levels and individual differences, several other variables were collected throughout the experiment. Participants' confidence in their decisions and understanding of the process were collected after each evaluation, whereas their overall effort levels were measured at the end of the experiment by using questions adapted from Davis (1989). In addition, La Trobe and Acott (2000) develop and validate a psychometric scale to assess one's environmental attitude. The scale consists of four dimensions to measure – (1) Human Interference with Nature; (2) Equity and Development Issues; (3) Humans and Economy over Nature; and (4) Duties to Nonhumans. To

[3] With respect to the GET framework, one may argue that including benchmark information would increase the knowledge that participants have when analyzing environmental performance measures, thus rendering the knowledge factor at a "high" level of evaluability for these measures (versus "low"). If the benchmark information had indeed led to high knowledge levels, then the evaluations in separate mode would also be made in a high evaluability setting since only one of the three factors need to be assessed as "high" for the information to be highly evaluable, and thus there would be no observed decision differences across evaluation mode. However, reported results later in the manuscript indicate that decision differences indeed occur across evaluation mode despite the inclusion of benchmark data. Thus, including benchmark data in the instrument enhances external validity while not introducing a confound to the experimental design.

[4] Willingness-to-invest responses are a type of willingness-to-pay measure. Such measures are prominent in marketing research (Breidert et al., 2006), and are also useful when considering environmental accounting judgments (Epstein, 2008).

simplify the task and avoid unnecessary fatigue, the study utilizes the third dimension to approximate participants' environmental attitudes. Finally, post-experimental questions collected data on demographic and academic status information.

2.5. Statistical methods

As suggested in Fig. 2, after evaluating environmental performance information in various mode and framing settings, each research participant ultimately makes judgments on eight factories, which are continuous but correlated measures. Traditional repeated-measures ANOVA assumes a multivariate normal distribution (MVN) and thus required either estimation of all covariate parameters of the MVN distribution or an assumption of sphericity of the covariate matrix (West et al., 2014). Instead, this study uses a Linear Mixed Model where the dependent variable is the evaluation decision made by each participant in four research groups, and the fixed effects are the evaluation mode and framing. Benchmark signal (positive or negative) is included as a covariate for the design reasons previously discussed. Task sequence and environmental attitudes are used as covariates to control for order effects and individual differences, respectively. Participant ID is used to define the residual covariance structure, whereas the evaluation task ID is used to indicate repeated observations. Mathematically, the LMM model can be expressed as:

$$\begin{aligned}\text{Evaluation Amounts} = {} & \beta_1 * \text{Evaluation Mode} && (\text{Joint vs.Separate; Fixed})\\ & + \beta_2 * \text{Framing} && (\text{Positive vs.Negative; Fixed})\\ & + \beta_3 * \text{Benchmark} && (\text{Better-or worse-than Benchmark; Fixed})\\ & + \beta_4 * \text{Task Order} && (\text{Fixed})\\ & + \beta_{5,6,7} * \text{Interactions of Mode, Framing, and Benchmark}\\ & + b_8 * \text{Environmental Attitude} && (\text{Random})\end{aligned}$$

where residual covariance is indicated by participant ID and repeated observation is indicated by task ID.

In summary, this study's experimental design includes a within-participant variable—the benchmark information valence (e.g., alternatives are both better/worse than the benchmark performance value). The overall result is a 2 (evaluation mode; between-participants) × 2 (performance measurement framing valence; between-participants) × 2 (benchmark information valence; within-participants) mixed factorial design.

3. Results

3.1. Research participants

Table 3 shows the descriptive statistics of the research participants. 54% of participants are female, 47% are undergraduate students, 50% are graduate students, 3% are non-degree-seeking students, and the average participant age is 28.97 years, indicating an older participant sample compared to many traditional student populations. To ensure successful random assignments of research participants, ANOVA and Chi-Square tests were performed on the four experimental groups, as well as mode (i.e., joint vs. separate) and framing (i.e., positive vs. negative) groups. The last columns of Table 3 show the resulting *p*-values, where no significant differences were detected on demographics and academic status.

ANOVA also indicates that research participants were statistically indifferent in their confidence in evaluation decisions ($p > 0.351$) and their understanding of the process ($p > 0.279$), suggesting between-group comparability. Also, comparisons of effort levels across experimental conditions suggest no difference in exerted effort ($p > 0.238$; see Table 3).

Table 3
Descriptive statistics.

	Overall	Groups				ANOVA *p*-values		
		JE Pos.	JE Neg.	SE Pos.	SE Neg.	Four-group	JE vs. SE	Pos. vs. Neg.
Age	28.97 (8.79)	29.56 (1.22)	28.12 (1.23)	27.84 (1.24)	30.28 (1.21)	0.439	0.837	0.681
Female	54.0%	51.9%	47.1%	60.0%	56.6%	0.589	0.210	0.571
Student type								
Undergraduate	97	23	24	23	27	0.869^	0.822^	0.757^
Graduate	102	29	24	24	25			
Non-degree	7	1	2	3	1			
Confidence	0 (1)	−0.05 (0.14)	−0.05 (0.14)	0.20 (0.14)	−0.07 (0.14)	0.490	0.432	0.351
Understanding	0 (1)	−0.09 (0.13)	0.25 (0.14)	0.01 (0.14)	−0.06 (0.13)	0.279	0.447	0.334
Effort	0 (1)	0.12 (0.14)	0.04 (0.14)	−0.02 (0.14)	−0.14 (0.14)	0.585	0.238	0.455

JE = joint evaluation mode; SE = separate evaluation mode.
^ denotes the p-values of a Pearson chi-square test. Due to non-degree-seeking students' low representation (less than 5), these numbers are combined with the graduate group.

3.2. Measurement validity

Table 4 shows the measurement validity of question items in this study. Both the Confidence and Understanding constructs are formative and synthesized with high factor loadings, with the average variance extracted (AVE) values being greater than the acceptance threshold (>0.51). Both the Effort and Environmental Attitude constructs are reflective and are collected in the post-experimental questions. The tests of their convergent and discriminant validities were satisfied, with the measurement items loading together and the square root of the AVE values are greater than the correlation between the constructs (−0.074, not tabulated; Fornell and Larcker, 1981).

3.3. Tests of hypotheses

Table 5 shows the tests of fixed effects by using a linear mixed model (LMM), where the dependent variable is the evaluation amount and independent variables are evaluation mode and measurement framing. Note that the impact of participants and the repeated observations (i.e., evaluation task) has been statistically controlled by LMM. In addition, benchmark, order effects, and environmental attitude (as a random effect) are mitigated as well. Due to the use of covariates in LMM, marginal estimated means (MEM) indicate the central tendency of the dependent variables in each experimental group *after controlling for the impact of control variables.*

Consistent with prior literature, participants give significantly higher evaluations when assessing environmental performance in joint than in separate mode (MEM = 54.69 vs.44.49), statistically significant at $p < 0.001$. Unsurprisingly, factories outperforming the benchmark receive better performance evaluations than those underperforming (MEM = 67.78 vs. 31.40). Further analyses on the significant interaction effect between mode and benchmark ($p < 0.001$) suggest that when the factory performance is worse than the benchmark, joint mode leads to significantly higher evaluations than separate mode (MEM = 39.85 vs. 22.95), but not when the factory performance is better than the benchmark. Such findings are consistent with similar reported

Table 4
Results of measurement validity.

Construct	Type	Items	Factor loadings	AVE^
Confidence	Formative	Example – How confident are you in your investment decision for Factory A? (1–7, 1 meaning "not at all," and 7 meaning "a lot.")	>0.676	0.53
Understanding	Formative	Example – How well do you understand the measurement, "Percentage of scrap materials not recycled back into production?" (1–7; 1 meaning "not at all," and 7 meaning "a lot.")	>0.737	0.60
Efforts	Reflective	I put a lot of effort into this task.	0.855	0.69
		I didn't try very hard to do well at this activity (Reverse)	−0.708	
		I tried very hard on this activity.	0.910	
Environmental attitude – humans and economy over nature	Reflective	Humans have the right to alter nature to satisfy wants and desires.	0.758	0.55
		Maintaining economic growth is more important than protecting the natural environment.	0.691	
		Humans have the right to reduce the number of species on Earth in order to promote economic development.	0.782	
		Humans do not have the right to subdue and control the rest of nature. (Reverse)	−0.744	

^ AVE (average variance extracted) shows the variance in the indicators or observed variables that is explained by the latent construct.

Table 5
Linear mixed model results.

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	810.538	4544.223	<0.001
Mode (i.e., joint vs. separate evaluation)	1	808.758	48.426	<0.001
Framing (i.e., positive vs. negative)	1	808.866	5.039	0.025
Benchmark (i.e., better- or worse-than benchmark)	1	808.551	618.019	<0.001
Task order	3	805.795	4.868	0.002
Mode * framing	1	809.593	9.139	0.003
Mode * benchmark	1	808.551	20.893	<0.001
Framing * benchmark	1	808.551	2.348	0.126

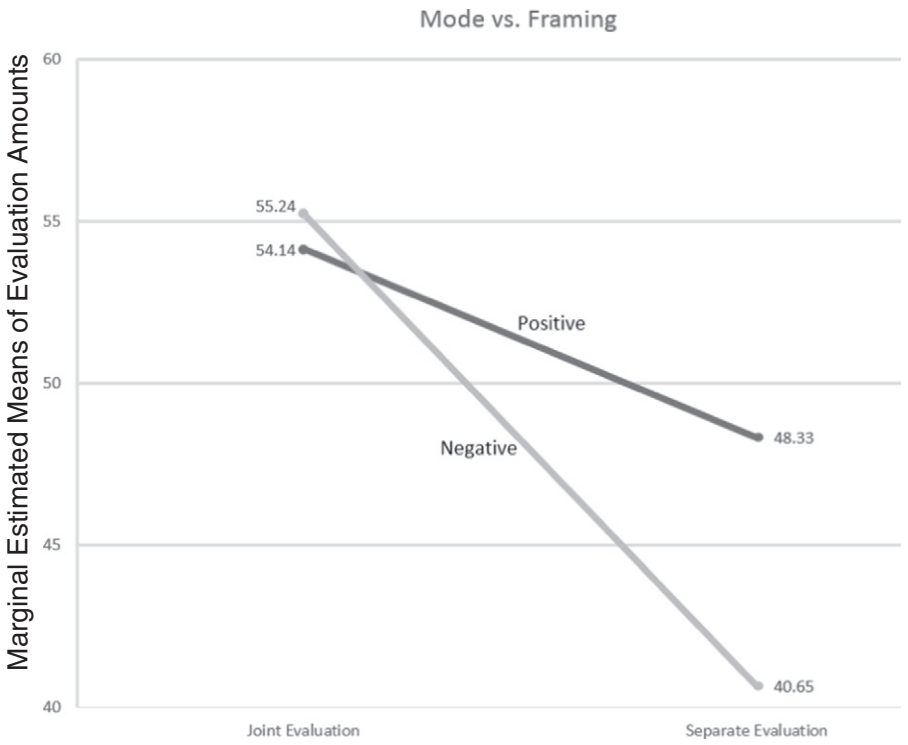

Fig. 3. Plot of evaluation mode and measurement framing effects on evaluations.

settings (Alewine and Stone, 2016). Probably due to the use of benchmark and the Latin-Square design, task order shows significant influences on evaluations ($p = 0.002$), which is statistically controlled in the model.[5]

H1 hypothesizes that higher (lower) evaluation judgments will occur when assessing environmental performance measures that contain positive (negative) framing. Evaluations of environmental performance measures containing positive framing were significantly higher compared to when the framing was negative (MEM = 51.24 vs. 47.95; $p = 0.025$). Results support H1 and suggest that the framing literature's observations of framing influences on judgments extend to decision settings involving environmental performance measures that contain characteristics of low evaluability.

The significant interaction of evaluation mode and measurement framing ($p = 0.003$) suggests the moderating effect of evaluation mode on framing valence's influence on environmental decisions. Specifically, Fig. 3 elaborates and visualizes the interaction effect and the MEM of evaluation mode and the measurement framing valence. When in separate mode, negative framing leads to significantly lower evaluations than positive framing does (MEM = 40.65 vs. 48.33; $p < 0.001$). However, statistical difference is not detected in joint mode when contrasting positive and negative framing (MEM = 54.14 vs. 55.24; $p = 0.595$). Also, evaluation differences across evaluation mode are relatively greater when framing is negative (MEM = 55.24 vs. 40.65; $p < 0.001$) compared to positive (MEM = 54.14 vs. 48.33; $p = 0.005$). Such results support H2–the performance measurement's framing effect on environmental judgments is moderated by the evaluation mode.

4. Discussion

This study extends the GET framework (Hsee and Zhang, 2010) to consider framing influences of performance information. Attribute framing appears to change how decision makers process knowledge of the performance information. In environmental accounting settings, measurement information is often not inherently evaluable, and users of the information contain low knowledge of the attributes. In such settings, GET suggests that the information's evaluation mode, joint or separate, will determine whether the information will ultimately be processed with high (joint) or low (separate) evaluability. The attribute framing's valence will influence measures more in separate mode, where the evaluability of the attributes are lower and thus more susceptible to being influenced by factors, such as attribute framing, whose decision weight will be more mitigated in joint evaluation mode due to more context in that mode in which to process the information.

This study also extends the attribute framing literature by, again, showing the consistency and robustness of attribute framing effects, but more importantly, by analyzing framing effects in a broad decision framework that allows for a general understanding

[5] Of the Latin Square Design's four tracks (see Table 2), two of the tracks started with a better-than-benchmark evaluation. These evaluations were lower than the two tracks that started with worse-than-benchmark evaluations.

of evaluability for attributes. Joint evaluation mode allows for greater contextualization of decisions, and our study suggests that this context serves to mitigate the impact of the attribute frames. When considering evaluability and contextualization, many other possible methods of contextualizing information could be considered. Multiple reference points (or benchmarks) should be investigated for a similar contextualization effect. Smith and Kida (1991) suggest that professional decision contexts can often mitigate bias, and our study may be showing such mitigation for attribute frames.

Practically, this study's reported findings advise AIS designers of cognitive influences to consider when designing AIS involving environmental information. Specifically, the environmental performance measures' framing influences decisions, and more so in separate than joint mode. This awareness of how attribute framing impacts low evaluability decisions is particularly important in environmental-related decisions, because while analyzing performance information in joint mode is intuitively more beneficial for effective judgments compared to in separate mode, many environmental decision settings exist where the joint presentation format is simply not an option due to limitations on alternate solutions. Thus, AIS designers would benefit from giving particular attention to attribute framing when generating reports for analysis in judgments. Based on our findings and past research on communication framing (Allport, 2010), negatively framing performance measures has a greater impact on decision makers when compared to positive framing. Our results suggest that while positively framed attributes in separate mode are significantly lower than in joint assessments, negatively framed attributes in separate mode are significantly lower than both joint assessments (regardless of the framing valence) and positively framed attributes in separate mode. This suggests that positively framed performance measures yield relatively more consistent judgments across evaluation mode than negative framing. In addition, AIS designers should seek to improve the evaluability of environmental accounting attributes via contextualization. If joint evaluation is not possible, additional benchmarks such as past performance may provide additional clarity to the decision setting.

Experimental designs sacrifice external validity by inducing simplistic approaches to what may be complex practitioner matters (see Shadish et al., 2002). This simplified approach to experimental design allows for a better understanding of causal mechanisms in place for how one variable impacts a judgment, and the environmental accounting literature has acknowledged the value of this approach (Holm and Rikhardsson, 2008; Kuruppu and Milne, 2010; Milne and Patten, 2002). Despite the literature's acceptance of this approach, this study attempts to mitigate any related limitations in at least two ways. First, strong cognitive psychological frameworks are employed to ensure that judgments do not occur as an artifact or demand effect of the experimental design or instrument. Second, external validity was enhanced by the careful implementation of real world factors such as adding benchmark information and utilizing realistic attribute descriptions. Care was taken to ensure that such factors did not drive the study's reported results.

Specific experimental decisions should also be mentioned as limiting the current study's findings. First, the decision was made to use only nonfinancial environmental attributes. The reasoning for this decision was environmental accounting information includes many nonfinancial aspects, and, as mentioned previously, qualitative studies and surveys have highlighted the importance and value of this nonfinancial environmental information (Albelda, 2011; KPMG, 2015; KPMG, 2011). Thus, a better understanding of the unique cognitive impact of processing such information is a necessary first step, and once the literature attains a better understanding of how these nonfinancial measures are processed in judgments, then future research can begin to enhance external validity of such judgments by including financial measures in decisions. Second, the study utilized positive and negative attribute frames consistently and exclusively across all decision attributes. One may certainly argue that a realistic decision setting would have different attribute frame types (positive and negative) and some non-framed attributes (strictly quantitative). However, our goal in this study was to analyze the robustness of the attribute framing effect in a controlled environmental accounting setting and to assess the ability of the joint evaluation mode to moderate this effect. To achieve these goals, the second in particular, the attribute framing effect needed to have appropriate power. It should also be said that past research has shown attribute framing effects with both different frame valences and non-framed attributes (Kuhberger, 1995; Kerler et al., 2015), and future researchers can investigate how these factors may impact the environmental accounting context.

Additional future research avenues abound for academicians who wish to explore behavioral issues relating to judgments involving environmental accounting information. How AIS will accommodate environmental accounting information will become more important for research considerations as managers increasingly require such information in decision settings. Overall, few environmental experimental studies exist (Alewine, 2010), yet experimental designs contain unique attributes and competitive advantages over other research methods that allow for the effective exploration of this type of accounting information, which contains uncommon and nontraditional characteristics that may result in unintended, yet unexplored, cognitive issues for processing such information. This study is an attempt to open our research colleagues to the rewarding possibilities of exploring such variables and research methods aimed at better understanding environmental AIS issues from a cognitive perspective and a decision making context.

5. Conclusion

This study analyzed the presentation of environmental accounting data in an AIS. We specifically analyzed the impact presentation mode (joint vs. separate evaluation) had on attribute framing. Study results suggest that attribute framing effects are present in decisions involving environmental attributes in separate evaluation mode, even with a base level of decision context/realism; positively (negatively) framed attributes yield higher (lower) evaluations. However, we find that in joint evaluation mode, the increased evaluability, through contextualization of the decision attributes, mitigates the effect of attribute frames on environmental judgments. Results inform environmental AIS designers and environmental accounting information users of cognitive consequences on judgments to consider with respect to evaluation mode and attribute framing factors in the AIS presentation design and information retrieval of environmental performance measures.

Appendix A. Research Instrument

A.1. Management decision making study

A.1.1. Introduction (all conditions)

Thank you for agreeing to participate in this study! On the following pages, please read the instructions and answer the questions.

This study will take <20 min to complete. When you are finished with the study, you will NOT be able to leave your seat until the administrator dismisses everybody at the same time. As a result, please take your time and focus when answering the questions. To thank you for your participation, you will have the opportunity to enter multiple random drawings to win a $25 gift card at the conclusion of the study.

In this study, you are a manager for Acme Company. One of your company's strategic objectives is to have good environmental business practices. Your job is to evaluate factories from the company and determine how well they are achieving this strategic objective.

The factories that you will evaluate are similar in size and produce the same items for the company. You will evaluate each factory's performance on two environmental measurements. Then, you will choose how much money you are willing to invest in each factory. **The better the factory's performance, the more you should invest in the factory**. The company believes that both measurements used to evaluate the factories are equally important in determining whether the factory achieves the strategic business objective.

You will be reminded of these instructions throughout the study.

Once you proceed to a new page, **please do not review prior pages**. As you proceed through the experiment, you should *not* base your new decisions on prior information or prior decisions.

Please turn the page when you are ready to proceed.

****Below shows a joint evaluation condition where performance measures contain a positive frame, and factory performances were better than the benchmark values.****

Reminder: you are a manager for Acme Company; one of your company's strategic objectives is to have good environmental business practices. Your job is to evaluate similarly constructed factories from the company and determine how well they are achieving this strategic objective. You will then choose how much money you are willing to invest in each factory. The company believes that both metrics used to evaluate the factories are equally important in determining whether the factory achieves the environmental strategic objective.

Please analyze data from two factories.

Metric	Factory A	Factory B
Percentage of scrap materials recycled back into production	83%	71%
Of the total compliments and complaints given by customers on environmental performance, the percentage of compliments	41%	48%

The more scrap materials recycled, the better the factory's performance.

The more customer compliments, the better the factory's performance.

You have been given the following information about similar factories:

- The average factory has scrap materials recycled back into production of 59%.

  Similar factories range from 30 to 89%.

- The average factory has customer compliments of 34%.

  Similar factories range from 17 to 51%.

For the questions below, please give an answer between 0 and 100, which represents how many millions of dollars you are willing to invest in each factory. *The better the factory's performance, the more you should invest in the factory*. These factories are *separate projects*; they are not competing with one another. You may invest up to 100 million in each factory. Please base your investment decision on how well you think the factories achieve the company's environmental strategic objective.

1. How much will you invest in Factory A? (0 − 100) _____
2. How much will you invest in Factory B? (0–100) _____

Please answer some questions about the evaluation you just completed.

For the questions below, please give an answer between 1 and 7, with 1 meaning “not at all,” and 7 meaning “a lot.”

1. How confident are you in your investment decision for Factory A? (1–7) _____
2. How confident are you in your investment decision for Factory B? (1–7) _____

For the questions below, please give an answer between 1 and 7, with 1 meaning "not at all," and 7 meaning "a lot."

1. How well do you understand the measurement, "Percentage of scrap materials recycled back into production?" (1–7) _____

2. How well do you understand the measurement, "Of the total compliments and complaints given by customers on environmental performance, the percentage of compliments?" (1–7) _____

Please do another evaluation. Your new evaluation should not be based on prior information or prior decisions!

****Below shows a separate evaluation condition where performance measures contain a positive frame, and the factory performance was better than the benchmark values.****

Reminder: you are a manager for Acme Company; one of your company's strategic objectives is to have good environmental business practices. Your job is to evaluate similarly constructed factories from the company and determine how well they are achieving this strategic objective. You will then choose how much money you are willing to invest in each factory. The company believes that both metrics used to evaluate the factories are equally important in determining whether the factory achieves the environmental strategic objective.

Please analyze data from the factory.

Metric	Factory
Percentage of scrap materials recycled back into production	83%
Of the total compliments and complaints given by customers on environmental performance, the percentage of compliments	41%

The more scrap materials recycled, the better the factory's performance.

The more customer compliments, the better the factory's performance.

You have been given the following information about similar factories:

- The average factory has scrap materials recycled back into production of 59%.

Similar factories range from 30 to 89%.

- The average factory has customer compliments of 34%.

Similar factories range from 17 to 51%.

For the question below, please give an answer between 0 and 100, which represents how many millions of dollars you are willing to invest in the factory. **The better the factory's performance, the more you should invest in the factory**. Please base your investment decision on how well you think the factory achieves the company's environmental strategic objective.

1. How much will you invest in the factory? (0–100) _____

Please answer some questions about the evaluation you just completed.

For the questions below, please give an answer between 1 and 7, with 1 meaning "not at all," and 7 meaning "a lot."

1. How confident are you in your investment decision for the factory? (1–7) _____

For the questions below, please give an answer between 1 and 7, with 1 meaning "not at all," and 7 meaning "a lot."

1. How well do you understand the measurement, "Percentage of scrap materials recycled back into production?" (1–7) _____

2. How well do you understand the measurement, "Of the total compliments and complaints given by customers on environmental performance, the percentage of compliments?" (1–7) _____

Please do another evaluation. Your new evaluation should not be based on prior information or prior decisions!

****After each round of evaluations, participants were told that the next evaluation should be made independently from all prior decisions. This process repeats until all factory pairings have been evaluated.****

****In the separate evaluation conditions, factories from the same pairing were not presented sequentially. This was done to help prevent comparisons across factories from the same pairing.****

****For all conditions, the following pages appear when the evaluations are complete.****

Thank you for your evaluations. This study is almost complete. Before the study concludes, please answer some basic questions about what you have done today.

For each of the following statements, please indicate how true it is for you, using a scale from 1 to 7, with 1 meaning "not true at all," 4 meaning "somewhat true," and 7 meaning "very true." There are no right or wrong answers. Please tell us how you really feel.

1	2	3	4	5	6	7
not at all true			somewhat true			very true

	Not at all true				Somewhat true		Very true
1. I have a good understanding of the task I just performed.	1	2	3	4	5	6	7
2. I put a lot of effort into this task.	1	2	3	4	5	6	7
3. In this task, the company was committed to protecting the environment.	1	2	3	4	5	6	7
4. I didn't try very hard to do well at this activity.	1	2	3	4	5	6	7
5. The task is too difficult to do effectively.	1	2	3	4	5	6	7
6. I understand the goals of doing this task.	1	2	3	4	5	6	7
7. I tried very hard on this activity.	1	2	3	4	5	6	7
8. This task helps me to achieve the company's goals.	1	2	3	4	5	6	7

For each of the following statements, please indicate what you think about the following statements, using a scale from 1 to 5, with 1 meaning “strongly disagree,” 3 meaning “not sure,” and 5 meaning “strongly agree.” There are no right or wrong answers. Please tell us how you really feel.

Please answer the following questions:

	Strongly disagree	Disagree	Not sure	Agree	Strongly agree
9. Humans have the right to alter nature to satisfy wants and desires.	1	2	3	4	5
10. There are limits to industrial growth.	1	2	3	4	5
11. Maintaining economic growth is more important than protecting the natural environment.	1	2	3	4	5
12. Natural resources should be used primarily to provide for basic needs rather than material wealth.	1	2	3	4	5
13. Humans have the right to reduce the number of species on Earth in order to promote economic development.	1	2	3	4	5
14. Humans have moral duties and obligations to other humans.	1	2	3	4	5
15. Humans do not have the right to subdue and control the rest of nature.	1	2	3	4	5
16. Present generations of humans have moral duties and obligations to future human generations.	1	2	3	4	5
17. Satisfaction and a high quality of life are less important than money or material wealth.	1	2	3	4	5

We really appreciate you taking the time to complete this study. There are only a few demographic questions remaining, and then you will be finished.

A.1.2. Demographic information

We need to ask you a series of demographic questions. Please answer the following questions:

Age: ____

Gender: ____

Are you an undergraduate, graduate, or non-degree seeking student? _____

Years of college education (include the current academic school year): ___

If you are pursuing a degree, indicate your subject emphasis. If this does not apply to you, select Not Applicable. Please circle one of the following: Not Applicable, Accounting, Economics, Finance, Management, Marketing, Other.

Amount of full-time work experience (round to the nearest year; enter “0” if not applicable): ___

If you have full-time work experience, please choose the field that best describes your type of experience. If this does not apply to you, select Not Applicable. Please circle one of the following: Not Applicable, Accounting, Economics, Engineering, Finance, Marketing, General Management, Engineering, Other

How many accounting classes are you currently taking? ___

How many accounting classes have you completed in the past (do not include classes you are currently taking)? ___

This concludes the study. Thank you for participating! After all participants have completed the study, we will have multiple drawings for a chance to win gift cards worth $25 each (limit one prize per participant). If you are a winner, then we will inform you by email. As you exit the room, you may sign up for these drawings.

Please remain in your seat until the administrator has dismissed all participants.

References

Abdallah, W.M., Alnamri, M., 2015. Nonfinancial performance measures and the BSC of multinational companies with multi-cultural environment. Cross Cult. Manag. 22 (4), 594–607.

Albelda, E., 2011. The role of management accounting practices as facilitators of the environmental management: evidence from EMAS organisations. Sustain. Account. Manag. Pol. J. 2 (1), 76–100.

Alewine, H.C., 2010. A model for conducting experimental environmental accounting research. Sustain. Account. Manag. Pol. J. 1 (2), 256–291.

Alewine, H.C., Stone, D., 2016. The joint influence of evaluation mode and benchmark signal on environmental accounting-relevant decisions. Soc. Environ. Account. J. 36 (2), 124–152.

Allport, C.D., 2010. Communication framing behavior and business information characteristics. J. Bus. Account. 3 (1), 147–159.

Ariely, D., Loewenstein, G., 2000. When does duration matter in judgment and decision making? J. Exp. Psychol. Gen. 129, 508–529.

Ariely, D., Loewenstein, G., Prelec, D., 2003. "Coherent arbitrariness": stable demand curves without stable preferences. Q. J. Econ. 118, 73–105.
Bacic, D., Appan, R., 2012. Impact of domain knowledge on information presentation format choice: a cognitive fit perspective. Proceedings of 2012 Americas Conference on Information Systems (AMCIS), Seattle, WA, pp. 26–36.
Ball, A., 2005. Environmental accounting and change in UK local government. Account. Audit. Account. J. 18 (3), 346–373.
Benston, G.J., 2006. Fair-value accounting: a cautionary tale from Enron. J. Account. Public Policy 25 (4), 465–484.
Blackburn, N., Brown, J., Dillard, J., Hooper, V., 2014. A dialogical framing of AIS–SEA design. Int. J. Account. Inf. Syst. 15, 83–101.
Bonner, S.E., 2008. Judgment and Decision Making in Accounting. Pearson Education, Inc., Upper Saddle River, New Jersey.
Breidert, C., Hahsler, M., Reutterer, T., 2006. A review of methods for measuring willingness-to-pay. Innovative Marketing 2 (4), 8–32.
Brockner, J., Wiesenfeld, B.M., Martin, C.L., 1995. Decision frame, procedural justice, and survivors' reactions to job layoffs. Organ. Behav. Hum. Decis. Process. 63 (1), 59–68.
Brown, D.L., Dillard, J.F., Marshall, R.S., 2005. Strategically informed, environmentally conscious information requirements for accounting information systems. J. Inf. Syst. 19 (2), 79–103.
Cater-Steel, A., Tan, W., 2010. The role of IT service management in green IT. Australasian J. Inf. Syst. 17 (1), 107–125.
Clarkson, P.M., Li, Y., Richardson, G.D., 2004. The market valuation of environmental capital expenditures by pulp and paper companies. Account. Rev. 79, 329–353.
Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. 13 (3), 319–340.
Debreceny, R.S., 2011. Betwixt and between? Bringing information systems and accounting systems research together. J. Inf. Syst. 25 (2), 1–9.
Dey, C., 2007. Social accounting at Traidcraft plc. Account. Audit. Account. J. 20 (3), 423–445.
Dikolli, S.S., Hofmann, C., Kulp, S.L., 2009. Interrelated performance measures, interactive effort, and incentive weights. J. Manag. Account. Res. 21, 125–149.
Dilla, W.N., Steinbart, P.J., 2005. Using information display characteristics to provide decision guidance in a choice task under conditions of strict uncertainty. J. Inf. Syst. 19 (2), 29–55.
Dilla, W., Janvrin, D.J., Raschke, R., 2010. Interactive data visualization: new directions for accounting information systems research. J. Inf. Syst. 24 (2), 1–37.
Dillard, J., 2008. Responding to expanding accountability regimes by re-presenting organizational context. Int. J. Account. Inf. Syst. 9, 21–42.
Dunegan, K., 1993. Framing, cognitive modes, and image theory: toward an understanding of a glass half full. J. Appl. Psychol. 78 (3), 491–503.
Epstein, M.J., 2008. Making Sustainability Work. Berrett-Koehler Publishers, Inc., San Francisco.
Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. J. Mark. Res. 18, 39–50.
Granlund, M., 2011. Extending AIS research to management accounting and control issues: a research note. Int. J. Account. Inf. Syst. 12, 3–19.
Gray, R., Bebbington, J., 2001. Accounting for the Environment. 2nd ed. SAGE Publications, London.
Herbohn, K., 2005. A full cost environmental accounting experiment. Acc. Organ. Soc. 30 (6), 519–536.
Holm, C., Rikhardsson, P., 2008. Experienced and novice investors: does environmental information influence investment allocation decisions. Eur. Account. Rev. 17 (3), 537–557.
Hsee, C.K., Leclerc, F., 1998. Will products look more attractive when presented separately or together? J. Consum. Res. 25 (2), 175–186.
Hsee, C.K., Zhang, J., 2004. Distinction bias: misprediction and mischoice due to joint evaluation. J. Pers. Soc. Psychol. 86 (5), 680–695.
Hsee, C.K., Zhang, J.A., 2010. General evaluability theory. Perspect. Psychol. Sci. 5, 343–355.
Hsee, C.K., Yang, Y., Li, N., Shen, L., 2009. Wealth, warmth, and well-being: whether happiness is relative or absolute depends on whether it is about money, acquisition, or consumption. J. Mark. Res. 46, 396–409.
Hsee, C.K., Zhang, J., Wang, L., Zhang, S., 2013. Magnitude, time, and risk differ similarly between joint and single evaluations. J. Consum. Res. 40 (1), 172–184.
Huang, A.H., 2009. A model for environmentally sustainable information systems development. J. Comput. Inf. Syst. 49 (4), 114–121.
Hughes, K.E., 2000. The value relevance of nonfinancial measures of air pollution in the electric utility industry. Account. Rev. 75, 209–228.
Humphreys, K.A., Trotman, K.T., 2011. The balanced scorecard: the effect of strategy information on performance evaluation judgments. J. Manag. Account. Res. 23, 81–98.
Jasch, C., Savage, D., 2008. The IFAC international guidance document for environmental management accounting. In: Schaltegger, S., Bennett, M., Burritt, R., Jasch, C. (Eds.), Environmental Management for Cleaner Production. Springer, New York, NY.
Jenkin, T., McShane, L., Webster, J., 2011a. Green information technologies and systems: employees' perceptions of organizational practices. Bus. and Soc. 50 (2), 266–314.
Jenkin, T., Webster, J., McShane, L., 2011b. An agenda for 'green' information technology and systems research. Inf. Organ. 21 (1), 17–40.
Joshi, S., Krishnan, R., Lave, L., 2001. Estimating the hidden costs of environmental regulation. Account. Rev. 76, 171–198.
Kaminski, J., 2003. Technologies and costs of SO_2-emissions reduction for the energy sector. Appl. Energy 75 (3–4), 165–172.
Kaplan, S.E., Wisner, P.S., 2009. The judgmental effects of management communications and a fifth balanced scorecard category on performance evaluation. Behav. Res. Account. 21 (2), 37–56.
Kelton, A.S., Pennington, R.R., Tuttle, B.M., 2010. The effects of information presentation format on judgment and decision making: a review of the information systems research. J. Inf. Syst. 24 (2), 79–105.
Kerler III, W.A., Allport, C.D., Fleming, A.S., 2012. Impact of framed information and project importance on capital budgeting decisions. Adv. Manag. Account. 21, 1–24.
Kerler III, W.A., Fleming, A.S., Allport, C.D., 2014. How framed information and justification impact capital budgeting decisions. Adv. Manag. Account. 23, 183–212.
Kerler III, W.A., Allport, C.D., Fleming, A.S., 2015. How framed information and justification impact capital budgeting decisions. J. Appl. Manag. Account. Res. 13 (1), 63–80.
Kogut, T., Ritov, I., 2005. The singularity effect of identified victims in separate and joint evaluations. Organ. Behav. Hum. Decis. Process. 97, 106–116.
KPMG, 2011. Corporate Sustainability: A Progress Report. KPMG.
KPMG, 2015. Currents of Change: The KPMG Survey of Corporate Responsibility Reporting 2015. KPMG.
Kuhberger, A., 1995. The framing of decisions: a new look at old problems. Organ. Behav. Hum. Decis. Process. 62 (2), 230–240.
Kuruppu, S., Milne, M.J., 2010. Dolphin deaths, organizational legitimacy and potential employees' reactions to assured environmental disclosures. Account. Forum 34 (1), 1–19.
La Trobe, H.L., Acott, T.G., 2000. A modified NEP/DSP environmental attitudes scale. J. Environ. Educ. 32 (1), 12–20.
Levin, I.P., Gaeth, G.J., 1988. Framing of attribute information before and after consuming the product. J. Consum. Res. 15 (3), 374–378.
Levin, I.P., Schneider, S.L., Gaeth, G.J., 1998. All frames are not created equal: a typology and critical analysis of framing effects. Organ. Behav. Hum. Decis. Process. 76 (2), 149–188.
Libby, R., Bloomfield, R., Nelson, M., 2002. Experimental research in financial accounting. Acc. Organ. Soc. 27, 775–810.
Lipe, M.G., Salterio, S.E., 2000. The balanced scorecard: judgmental effects of common and unique performance measures. Account. Rev. 75 (3), 283–298.
Mălăescu, I., Sutton, S.G., 2015. The effects of decision aid structural restrictiveness on cognitive load, perceived usefulness, and reuse intentions. Int. J. Account. Inf. Syst. 17, 16–36.
Marelli, A., 2015. The evolving role of environmental management accounting in internal decision-making: a research note. Int. J. Account. Audit. Perform. Eval. 11 (1), 14–47.
Milne, M.J., Patten, D.M., 2002. Securing organizational legitimacy: an experimental decision case examining the impact of environmental disclosures. Account. Audit. Account. J. 15 (3), 372–405.
Morewedge, C.K., Kassam, K.S., Hsee, C.K., Caruso, E.M., 2009. Duration sensitivity depends on stimulus familiarity. J. Exp. Psychol. Gen. 138, 177–186.
O'Donnell, E., David, J.S., 2000. How information systems influence user decisions: a research framework and literature review. Int. J. Account. Inf. Syst. 1, 178–203.
Rom, A., Rohde, C., 2007. Management accounting and integrated information systems: a literature review. Int. J. Account. Inf. Syst. 8, 40–68.
Romney, M.B., Steinbart, P.J., 2015. Accounting Information Systems. 13th ed. Pearson Education, Inc., Upper Saddle River, NJ.
Schaltegger, S., Burritt, R., 2010. Sustainability accounting for companies: catchphrase or decision support for business leaders? J. World Bus. 45, 375–384.
Schweitzer, M., 1995. Multiple reference points, framing, and the status quo bias in health care financing decisions. Organ. Behav. Hum. Decis. Process. 63 (1), 69–72.
Shadish, W.R., Cook, T.D., Campbell, D.T., 2002. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton Mifflin, Boston.

Shaft, T.M., Vessey, I., 2006. The role of cognitive fit in the relationship between software comprehension and modification. MIS Q. 30 (1), 29–55.
Simnett, R., Vanstraelen, A., Chua, W.F., 2009. Assurance on sustainability reports: an international comparison. Account. Rev. 84 (3), 937–967.
Smith, M.J., 2002. Gaming nonfinancial performance measures. J. Manag. Account. Res. 14, 119–133.
Smith, J.F., Kida, T., 1991. Heuristics and biases: expertise and task realism in auditing. Psychol. Bull. 109 (3), 472–489.
Sriramachari, S., Chandra, H., 1997. The lessons of Bhopal [toxic] MIC gas disaster scope for expanding global biomonitoring and environmental specimen banking. Chemosphere 34 (9–10), 2237–2250.
Stone, D., Schkade, D., 1994. Effects of attribute scaling on process and performance in multiattribute choice. Organ. Behav. Hum. Decis. Process. 59 (2), 261–287.
Sullivan, K., Kida, T., 1995. The effect of multiple reference points and prior gains and losses on managers' risky decision making. Organ. Behav. Hum. Decis. Process. 64 (1), 76–83.
Vasarhelyi, M.A., 2012. AIS in a more rapidly evolving era. J. Inf. Syst. 26 (1), 1–5.
Vessey, I., 1991. Cognitive fit: a theory-based analysis of the graphs versus tables literature. Decis. Sci. 22 (2), 219–240.
Weigand, H., Elsas, P., 2012. Construction and use of environmental management accounting systems with the REA business ontology. J. Emerging Technol. Account. 9 (1), 25–46.
West, B.T., Welch, K.B., Galecki, A.T., 2014. Linear Mixed Models: A Practical Guide Using Statistical Software. CRC Press.
Wines, G., Dagwell, R., Windsor, C., 2007. Implications of the IFRS goodwill accounting treatment. Manag. Audit. J. 22 (9), 862–880.
Yeung, C., Soman, D., 2007. The duration heuristic. J. Consum. Res. 34, 315–326.
Zhang, Y., Buda, R., 1999. Moderating effects of need for cognition on responses to positively versus negatively framed advertising messages. J. Advert. 28 (2), 1–15.