The meaning of a defined accounting concept: regulatory changes and the effect on auditor decision making

Jane J.F. Hronsky *, Keith A. Houghton

Department of Accounting, Faculty of Economics and Commerce, University of Melbourne, Parkville, VIC, Australia

Abstract

Accounting standards exist in an attempt to “standardize” accounting practice. These standards contain definitions of accounting concepts whose function is to guide judgments made in practice. However, such judgments can have a major impact on a firm’s externally reported accounting numbers, as their inherent subjectivity and discretion may be lent to the manipulation of earnings. This study provides empirical evidence of the effect of measured meaning on an accounting judgment, in the context of regulated changes to the definition of one key accounting concept used in measuring operating income. The extraordinary items classification decisions made by auditors were found to be systematically associated with differences in measured meaning of the extraordinary items definition. The study has important policy implications for accounting standard-setting.

1. Introduction

A firm’s external financial reports are the end product of numerous judgments and decisions. In making these decisions, some authoritative guidance is provided by accounting standards. However, rule-based systems of regulation can never be ‘watertight’, and allegations of ‘creative’ or ‘aggressive’ accounting practices by firms are often heard (for example, Griffiths, 1986; Smith, 1992). Creative accounting can be defined as “the process by which management takes advantage of gaps or ambiguities in the standards to present a biased picture of financial performance. It does not breach the letter of the law or rules, but may breach its spirit” (Shah, 1998, p. 83). An aggressive reporting environment has been characterized as an ambiguous one in which definitive applicable authoritative guidance is lacking (Kennedy, Kleinmuntz & Peecher, 1997; Phillips, 1999). Some have suggested (e.g. Shah, 1996) that the practice of ‘creative compliance’ makes standard-setting appear to be a redundant exercise.

Nevertheless, it is possible in the language game of standard-setting, for regulators to use language carefully to close ‘gaps’ in the rules and amend vague and/or incomplete rules.¹ Mason and Gibbins (1991, p. 23) suggest that it would be useful to

¹ This is not to deny, however, that standard-setting often appears to exhibit what Shah (1996) refers to as a “dialectic of creativity”; that is, from avoidance to rules to avoidance again; a game of “catch-up” for regulators.
“Review new (and possibly existing) standards to reduce the large number of ambiguities and other difficulties that detract from the thrust of the standard by requiring interpretative and clarifying judgments rather than judgments on matters of substance”; and “State the main objectives of each accounting standard, to facilitate the task of statement preparers and auditors who are trying to apply the standard according to its spirit, rather than according to a legalistic interpretation of its words”.

Importantly, clearly-worded standards provide guidance for auditors in auditor–client conflict situations, and reduce the justifiability of aggressive reporting decisions. Auditors are a major influence on the form of the final published financial statements; the statements can be viewed as effectively a joint product of management and the auditor (Gibbins, Richardson & Waterhouse, 1990; Gibbins, Salterio & Webb, 1998; Antle & Nalebuff, 1991; Hansen & Watts, 1997); with the external auditor spending up to 30% of their time in financial statement preparation and finalization (Hackenbrack & Knechel, 1997). Knapp (1985) found that financial users perceived that management was more likely to obtain its preferred resolution in a conflict with auditors when the conflict issue was not dealt with precisely by technical standards, as opposed to the situation where the conflict issue was dealt with precisely by technical standards. He also found (Knapp, 1987) that audit committee members were more likely to support an auditor involved in a major dispute with management when the focal issue was the subject of objective technical standards.

This study examines one instance of regulators playing, apparently successfully, the “language game”: the change in Australian accounting standards of the definition of an “extraordinary item”. The intention of the regulators was to remove the flexibility inherent in the existing definition, and thus limit the inconsistencies and alleged opportunism observed in practice. The study finds evidence that auditors do ascribe different meanings to the old and new definitions of an extraordinary item, and that those differences are systematically associated with different extraordinary item classification decisions.

In doing so, the study also builds upon the extant measurement of meaning in accounting literature. That literature has typically been concerned with the extent to which meanings of basic accounting terminology are shared (or not shared) between the parties to the communication process (e.g. between preparers and users). Extending the analysis to the effect of meaning on behavioral outcomes establishes a link between this body of literature and the much larger literature dealing with judgment in accounting and auditing.

This study seeks to determine whether decision outcomes are a function of, inter alia, the meanings attributed to the accounting signal (the definition of extraordinary item contained in the accounting standard) on which the decisions are based. The study finds this relationship to exist, and as such, has important policy implications for accounting standard setting and other regulation. The methodology employed provides a means by which regulators may prospectively evaluate the impact of changes in meaning of accounting concepts and other words used to regulate accounting numbers and disclosures on accounting policy decisions.

The paper is organized as follows. Section 2 discusses the relevant prior literature and formulates the hypotheses. Section 3 describes the research methodology. Section 4 reports the data analysis and results, while Section 5 discusses those results. Section 6 provides a summary and conclusion.

2 The study of meaning, language and communication has been approached from various theoretical viewpoints, including those developed in philosophy, semiotics, linguistics, psychology and media studies. Although different emphases may be given in alternative models of communication, certain elements exist in common: a sender/writer/producer, a receiver/reader/user, a channel/medium, a message/text and an effect. In this study, the particular emphasis is on the reader (auditors), the message (the accounting standard definition) and the effect (the decisions made on the basis of that definition). We are interested in the interpretation; in the process of assigning meanings to the message.

3 ‘Signal’ is used in the sense of the physical existence or form of the text; that is, the formal definition of ‘extraordinary item’ as printed in the applicable accounting standard.
2. Prior literature and hypotheses

2.1. The extraordinary items judgment

Accounting regulation in many parts of the world (including Australia) requires the separation of extraordinary items from operating income. The accounting policy decision to classify an event or transaction as ‘extraordinary’ is made if it falls within the definition of ‘extraordinary’ provided by the relevant accounting standard. Until recently the treatment of extraordinary items (EIs) in Australia has been seen as a method of profit manipulation. Serious concern was expressed (see, for example, Jukes, 1988; Parry, 1990; Rennie, 1985; Walker, 1988) about the flexibility inherent in the relevant definition:

Extraordinary items means items of revenue and expense and other gains and losses ... attributable to events or transactions outside the ordinary operations of the business entity. [AAS 1, 1974, para 4(b)].

In particular, criticism was levelled at the exclusion of extraordinary loss items from net operating profit, while gains of a similar nature were included in operating profit.4 In an attempt to overcome this problem, the standard was amended and reissued. The crucial change to the definition was the addition of the words “and not of a recurring nature”. The intention of the small but important change in wording was to limit the extraordinary classification to a few isolated cases and end the inconsistencies (and arguably the opportunistic behavior) that had been evident in practice.

The inclusion or exclusion of items from operating profit to manipulate the accounting income number has been empirically observed over a number of years. A comparison of the studies is difficult due to their differing methodologies, objectives and times of execution. However, a number of general conclusions may be drawn from the literature. The evidence suggests that EIs are used as discretionary adjustments to income, consistent with both the ‘big bath’ and ‘income smoothing’ hypotheses (for example, Barnea, Ronen & Sadan, 1975; Craig & Walsh, 1989; Godfrey & Jones, 1999; Walsh, Craig & Clarke, 1991). Hoffman and Zimmer (1994) found that, in 1989, the tendency to classify recurring gains rather than losses as operating instead of extraordinary was greater where the chief executive received a higher remuneration relative to firm earnings. Recent capital market studies (e.g. Easton, 1990) suggest that EIs have little or no information content in assessing security values. Houghton and Walker’s (1991) experimental study examined the impact of changes to the regulatory environment in which the extraordinary/operating classification decision is made in Australia, and found evidence to suggest that the change may have had the desired effect of reducing the manipulation of income through the categorization of EIs.

This study extends previous research by determining the extent to which the perceived change in the meaning of the definition of an EI contained in the relevant accounting standard, is driving the classification decisions made on the basis of those definitions.

2.2. Measurement of meaning in accounting

As noted above, by focusing on the effect of a change in the definition of an accounting concept, this study also extends the measurement of meaning in accounting literature. This literature has largely applied the semantic differential methodology originally developed by Osgood, Suci and Tannenbaum (1957) in experimental psychology, to the accounting domain of meaning. The semantic differential instrument provides a

---

4 The following comment is typical of the concerns expressed by financial commentators. “The extraordinary items classification is just too subjective ... the more cynical commentators would have us believe that the extraordinary item line has become a receptacle for large debits, while credits which might even fit the bill as extraordinary have somehow ended up as part of operating profits” (Jukes, 1988, p. 85).
Studies conducted to date in the accounting domain have primarily focussed on the extent to which various parties to the communication process attribute the same meanings to key accounting concepts. This literature has been critically reviewed by Bagrano ('90). Most studies have looked at the differential meanings of basic accounting concepts held by preparers and users of financial accounting information (Haried, '73; Houghton, '87a, '88; Houghton & Messier, '90; Karvel, '79; Oliver, '74), managerial accounting information (Johnson, '92), and, more recently, the differential meanings held by auditors across national cultural boundaries (Bagrano, Houghton & Hronsky, '94). Studies with an educational focus have shown that the meanings of accounting concepts held by students change over time (Houghton, '87b) and that the meanings held by inexperienced members of the accounting profession are not identical to those held by practicing accountants (Houghton & Hronsky, '93). Importantly, it has been shown that the dimensions of the accounting domain of meaning conform to the Evaluative–Potency–Activity dimensions of the general domain of meaning (Houghton, '88) and that subjects with a level of knowledge or sophistication about accounting information exhibit this degree of cognitive complexity (Houghton, '87a, '88).

To date, this literature does not address the question of what, if any, differences in the perceived meanings of accounting concepts held by individuals have on accounting practice. In the context of the EI classification decision, this study addresses the impact of connotative meaning on accounting decision making.

2.3. Research question and hypotheses

The fundamental research question of this study is to determine whether there is a measurable and empirically verifiable link between decision behaviors (outcomes) and connotative or measured meaning: whether some of the variability in behavioral responses can be explained by variability in the measured meaning of the decision rule (the EI definition). Variability in the meaning of the EI concept is likely to be present given the recent change in its definition in the accounting standards. The first hypothesis to be tested is therefore concerned with whether this variability in meaning can be established.

\[ H_0 \]. There is no significant difference between the meaning of the concept 'extraordinary item' in the old Australian definition and the new definition.

The next step involves examining the decisions made on the basis of the concept. As a classification decision must be either 'extraordinary' or 'operating' (the two are mutually exclusive) the

---

5 Osgood et al. (1957) rely on the distinction made by many (from John Stuart Mill onwards) between denotative and connotative meaning. Denotative meaning is the ordinary or literal meaning of a concept; connotative meaning is a subjective or emotional meaning. Osgood et al. (1957, p. 323) observed that even with agreement between communicating parties as to denotative meaning the parties can display different behaviours in their response to a concept; thus implying the importance of connotative meaning in driving individual behaviour. Section 6 contains a discussion of the potential shortcomings associated with the assumption that denotative or literal meanings are shared, and that the important variation in constructed meaning takes place in the realm of the connotative.

6 To quantify their meaning of a concept, subjects mark the space between pairs of adjectives which best describes their association between the descriptive adjectives and the concepts. The meaning of a concept is determined by its location within 'semantic space'. Semantic space is multi-dimensional Euclidean geometrical space, with each dimension representing an independent dimension of meaning. Osgood et al. (1957) found the existence of three dimensions: the Evaluative dimension, (represented by the semantic differential scale ‘good–bad’); Potency (‘strong–weak’); and Activity (‘active–passive’). The meaning of a concept is measured by the placement of that concept on the axes of these dimensions. The existence of the three factor E–P–A structure has been validated in numerous studies (Heise, 1969) with the psychological research finding these three factors regardless of the domain of interest. The methodology has been, and continues to be, widely used in numerous disciplines, including psychology (Goldsmith, McDermott & Alexander, 2000), sociology (Tzeng & Henderson, 1999) and education (Tracey, Aroll & Richardson, 1999).
necessary variability in decision outcomes is present. The hypothesis may be stated in the null form as:

\[ H_0 \]
There is no significant difference between the classification decisions made by subjects on the basis of the old Australian definition and on the basis of the new definition.

If these two hypotheses are able to be rejected, it means that there is a significant amount of variability in the meaning of the accounting concept; and a significant amount of variability in the decision outcomes based on this concept. That these two conditions be present is necessary (but not sufficient) to establish a link between the two. The linkage is the fundamental research question of this study, and may be stated as:

\[ H_3 \]
That variability in extraordinary classification decisions may be explained by the variability in the meaning of the concept upon which decisions are based.

If a significant association between measured meaning and decision outcomes is found, it will provide evidence in understanding how changes to accounting standards and regulations flow through to impact on decisions made by those involved in the preparation of accounting information. It will also empirically establish a linkage between connotative meaning and an accounting decision, and provide some explanation of the role of meaning in accounting decision making.

3. Research methodology

3.1. Research design

This study used a between-subjects design in respect of the old and new regulated definitions. The group designated as Group 1 (NEW) were to be prompted by the new Australian definition of EI in making their decisions, while Group 2 (OLD) were to be prompted by the old definition. Each of the groups were to be presented with the test instrument requiring them to:

1. make a decision as to whether the facts about particular events or transactions would result in them being classified as either ‘extraordinary’ or ‘operating’ for each of 10 cases (the two categories being mutually exclusive); and then
2. complete a semantic differential to ascertain their meaning of the concept of EI used in making these decisions.

The test instrument consisted of two parts to be completed by each subject; with the first part designed to test for variability in the classification decisions, and the second part designed to test the meanings of the concepts used by subjects. An extract of a test instrument is included as an Appendix.

3.2. Case selection

Part One of the instrument contained 10 cases, all somewhat simplified, based on real but disguised Australian examples of items where the exercise of judgment (classification as an EI or not) was necessary. All items were material. The

---

7 Two responses were elicited from subjects in respect of the classification decision. Requirement 1 was the dichotomous classification (extraordinary/operating). Requirement 2 was the same classification over a range of 1 to 6, from ‘clearly extraordinary’ to ‘clearly operating’. These two decisions, as may be expected, are highly correlated (Spearman coefficient = 0.8016, \( P < 0.001 \)) and may therefore be used interchangeably. Responses from Requirement 2 were used in the data analysis as they have the better distributional characteristics.

8 As noted above, this study measures only connotative meanings. The research design controls for possible variation in denotative meanings by the following means: subjects all had expertise in the specialized discourse of accounting practice (all had received in-house technical training about the new standard); were from a shared geographical/cultural space at the same point in time; and were randomly assigned to the old and new definition groups.

9 Six of the cases were based on those used in Houghton and Walker (1991) and are used with the authors’ permission; an additional four cases were developed based on information obtained from the annual reports of listed Australian companies.
cases are therefore representative of the types of events and transactions likely to be encountered in practice. The cases include both recurring and non-recurring items and also both gains and losses to test for differential treatment. Subjects were required to designate each case as representing either an ‘extraordinary’ or ‘operating’ item. These two alternatives are mutually exclusive and form the comprehensive list of available options.

3.3. Semantic differential

All studies of the measurement of meaning in accounting have drawn (in whole or part) on scales developed in the original work of Osgood et al. (1957) or those developed by Haried (1972), to represent independent dimensions of meaning (factors). Consistent with Haried (1973) and others, this study uses a sub-set of Haried’s original 33 scales. A factor loading of greater than 0.5 was used as the selection criterion. The resulting 22 scales may be seen in Table 2 and the Appendix. The scales have been used and validated in a number of studies (Bagranoff et al., 1994; Houghton 1987a,b; Houghton & Hronsky, 1993; Houghton & Messier); and are accordingly accepted as valid for the present study. The 22 scales were randomized in order with the right and left sides randomly “flipped”.

In semantic differential tasks, subjects are typically required to mark one of the seven spaces between the semantic differential scales to indicate their perceived relationship between the concept EI and the descriptive scales. A restatement of the definition applied by the subject in the preceding judgment task (the classification of the 10 cases) was given at the head of the page. Subjects were instructed to evaluate the concept EI in the context of that definition.

3.4. Subject selection

Experienced auditors were used as the subjects of this study. Although they may not make the initial classification decision, in practice they review many such decisions made in different circumstances, and have significant input to the final nature and amount of disclosures made in the financial statements. They therefore have a broad exposure to those types of decisions, as well as being more easily accessible as experimental subjects. An experienced auditor was defined as one with at least 3 years’ audit experience. The final sample contained 20 subjects per group. Subjects were voluntary participants from four of the Big 6 accounting firms in one of the two large commercial centres in Australia (Melbourne). The firms were KPMG (Peat Marwick); Arthur Andersen & Co; and the then Coopers & Lybrand and Price Waterhouse.

3.5. Administration

The data was collected under laboratory conditions in each of the offices of the relevant accounting firms. The occasion for the collection was training sessions run for audit staff during normal working hours. The data was collected in seven sessions over a 7 week period. At the time of the instrument administration, the revised definition of EI had been in operation for almost a

---

10 The distribution of cases was 50% gains and 50% losses, and 70% recurring items and 30% non-recurring items. Although the sample cases do not contain equal proportions of both of these variables, they are indicative and representative of real-world events and transactions in the area. Given that the focus of the change in definition was on the elimination of recurring EI, it makes sense to include a greater number of recurring cases.

11 The sole exception is Karvel (1979), who determined for himself some additional scales.

12 The common technique for assigning scales to factors is to select those which load most heavily, which is defined conventionally as greater than 0.5. See, for example, Houghton (1988).

13 Any subjects with less than 3 years’ experience were deleted from the sample. The initial sample contained some subjects who were rejected as they had less than the required (3 years) experience. The mean years of audit experience was 5.10 with a standard deviation of 1.15 years. Three years was seen as sufficient experience to make an initial judgment, although it must be borne in mind that these judgments may be subject to review in practice.
Subjects had therefore been through training sessions and had experience in using the new definition. One or both of the present writers were present at all sessions to ensure conformity of instructions and subject conditions. Subjects were randomly assigned to treatment groups. Subjects took approximately 30 minutes to complete the instrument; debriefing suggested no subject fatigue was apparent. All instruments were completed except for a single case decision for a single subject.

4. Results

The data analysis is divided into three parts, following the structure of the three hypotheses. The first part deals with the analysis of the semantic differential data to determine whether there are significant between-group differences in meaning of the concept EI. The second part of the analysis deals with the analysis of the extraordinary/operating classification decisions and whether there are significant between-group differences in decision outcomes. The final part deals with the association between measured meaning and decision outcomes.

4.1. Factor analysis of semantic scales

Consistent with the existing measurement of meaning literature, the data reduction process used for the semantic differential data is factor analysis with varimax (orthogonal) rotation (see for example, Houghton, 1987a,b; Osgood et al., 1957). Factor analysis of all responses showed six factors with an Eigenvalue of greater than one, as shown in Table 1.

It can be seen that more than half of the variance in responses is explained by the first three factors. Applying both the simple scree test and the more rigorous factor comparability test (Everett & Entrekin, 1980) resulted in the adoption of a robust and stable three factor solution. To determine if all subjects shared the same nature of factors (that is, consisted of similarly highly correlated scales), as well as the same number, between-group tests of factor comparability were made. Individual factor structures were highly comparable; that is, subjects shared the same complex cognitive structure for accounting concepts. The analysis therefore proceeds with the overall factor structure which is shown in Table 2.

Fig. 1 shows the nature of the factors obtained. It shows the semantic differential scales that load heavily on, or correlate highly with, each of the three factors, in decreasing order of loading.

The three factors show consistency with the standard Evaluative–Potency–Activity structure first identified by Osgood et al. (1957), although not in that order. The first factor is the most straightforward to describe. All scales have some idea or aspect of potency or substantiveness about them; this factor is accordingly labelled POTENCY.

Most of the scales loading heavily on the third factor relate in some way to the notion of activity: variable, short term, dynamic, temporary. However, some of the most heavily loading of the scales appear to relate to the notion of ‘controllability’ or ‘manageability’ of that activity: unexpected–expected, planned–unplanned and controllable–uncontrollable. As the nature of a possible EI in terms of ‘activity’ (whether it is recurring or non-recurring) is important to its classification, it is

---

14 It can be argued that, given that timing of the administration was after the introduction of the new standard, subjects in Group 2 (the old definition of EI) would not fully incorporate the old regulatory position into their decision making. That is, the timing of the administration would work against finding a significant difference between the decisions of Group 1 and Group 2. As shown below, if this bias exists, it was not present in sufficient force to contaminate the results to that extent. A different problem is the possibility that until the new benchmarks have been internalized by auditors, decision patterns may be somewhat erratic. Mitigating against this is the fact that in practice difficult decisions would be subject to review by a superior.

15 The total population was randomly split into two halves and their factor scores compared by way of a Pearson correlation. The coefficients obtained across the factors were 0.9865, 0.9976 and 0.9776, all highly significant \( (P < 0.001) \). The three factor solution is therefore stable and robust.

16 A separate factor analysis was performed for each of the subject groups, and correlations between their factor scores calculated. These were 0.9835, 0.9782 and 0.9473, all highly significant \( (P < 0.001) \). These highly comparable factor structures indicate the factor structure of the group was shared.
interesting to note this idea of the planned or expected nature of that item is incorporated into the dimension of the cognitive structure which is labelled ACTIVITY.

The second factor is the least clear-cut to describe. The scales beneficial–adverse and bad–good are clearly evaluative; however, the other scales which load on this factor would intuitively seem to be more related to factor one. It is possible that its perceived substantiveness or potency is relevant for the evaluation of the EI definition. As the evaluative scales are the most heavily loading, this factor is labelled EVALUATIVE.

Thus, the factor structure may be described as follows:

- Factor 1: POTENCY
- Factor 2: EVALUATIVE
- Factor 3: ACTIVITY

This factor structure might be thought of as a three dimensional structure (a cube) within which the concepts (the definitions) are held. The meanings of concepts are determined by their relative placements on the factor structure (or within the cube). Table 3 shows the placements of the concept EI on the factor structure by the different subject groups.

To determine if there are significant between-group differences in meaning, analysis of variance (ANOVA) was performed on the factor scores by group.

Table 1
Eigenvalues and proportion of variance explained—all groups

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>% of variance</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.7626</td>
<td>30.7</td>
<td>30.7</td>
</tr>
<tr>
<td>2</td>
<td>2.2807</td>
<td>10.4</td>
<td>41.1</td>
</tr>
<tr>
<td>3</td>
<td>2.2163</td>
<td>10.1</td>
<td>51.2</td>
</tr>
<tr>
<td>4</td>
<td>1.6881</td>
<td>7.7</td>
<td>58.9</td>
</tr>
<tr>
<td>5</td>
<td>1.3395</td>
<td>6.1</td>
<td>64.9</td>
</tr>
<tr>
<td>6</td>
<td>1.0263</td>
<td>4.7</td>
<td>69.6</td>
</tr>
<tr>
<td>7</td>
<td>0.8806</td>
<td>4.0</td>
<td>73.6</td>
</tr>
</tbody>
</table>

Table 2
Varimax rotated factor matrix: three factor solution for all treatment groups

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Exact–estimated</td>
<td>0.86153(^a)</td>
<td>0.18253</td>
</tr>
<tr>
<td>2</td>
<td>Bad–good</td>
<td>0.24757</td>
<td>0.70069(^a)</td>
</tr>
<tr>
<td>3</td>
<td>Measurable–unmeasurable</td>
<td>0.57251(^a)</td>
<td>0.47671</td>
</tr>
<tr>
<td>4</td>
<td>Necessary–unnecessary</td>
<td>0.13669</td>
<td>0.64958(^a)</td>
</tr>
<tr>
<td>5</td>
<td>Planned–unplanned</td>
<td>0.16991</td>
<td>0.28592</td>
</tr>
<tr>
<td>6</td>
<td>Objective–subjective</td>
<td>0.75514(^a)</td>
<td>0.01677</td>
</tr>
<tr>
<td>7</td>
<td>Tangible–intangible</td>
<td>0.53898(^a)</td>
<td>0.28512</td>
</tr>
<tr>
<td>8</td>
<td>Strong–weak</td>
<td>0.26083</td>
<td>0.60724(^a)</td>
</tr>
<tr>
<td>9</td>
<td>Indirect–direct</td>
<td>0.35607</td>
<td>0.29609</td>
</tr>
<tr>
<td>10</td>
<td>Variable–constant</td>
<td>0.30247</td>
<td>0.10547</td>
</tr>
<tr>
<td>11</td>
<td>Safe–risky</td>
<td>0.57494(^a)</td>
<td>0.32796</td>
</tr>
<tr>
<td>12</td>
<td>Complete–incomplete</td>
<td>0.76653(^a)</td>
<td>0.22056</td>
</tr>
<tr>
<td>13</td>
<td>Discretionary–required</td>
<td>0.41604</td>
<td>0.30691</td>
</tr>
<tr>
<td>14</td>
<td>Real–imaginary</td>
<td>0.39571</td>
<td>0.61816(^a)</td>
</tr>
<tr>
<td>15</td>
<td>Beneficial–adverse</td>
<td>0.04889</td>
<td>0.74985(^a)</td>
</tr>
<tr>
<td>16</td>
<td>Temporary–permanent</td>
<td>0.10716</td>
<td>0.13900</td>
</tr>
<tr>
<td>17</td>
<td>Controllable–uncontrollable</td>
<td>0.06313</td>
<td>0.27639</td>
</tr>
<tr>
<td>18</td>
<td>Unexpected–expected</td>
<td>0.00367</td>
<td>0.21473</td>
</tr>
<tr>
<td>19</td>
<td>Passive–active</td>
<td>0.08477</td>
<td>0.49117</td>
</tr>
<tr>
<td>20</td>
<td>Static–dynamic</td>
<td>0.08118</td>
<td>0.17602</td>
</tr>
<tr>
<td>21</td>
<td>Short term–long term</td>
<td>0.16688</td>
<td>0.13202</td>
</tr>
<tr>
<td>22</td>
<td>Inflexible–flexible</td>
<td>0.69399(^a)</td>
<td>0.38492</td>
</tr>
</tbody>
</table>

\(^a\) Loadings of greater than 0.5.
A highly significant difference in placement was found on Factor 1 (POTENCY) between Group 1 (NEW) and Group 2 (OLD) \( (F = 116.88, P < 0.001) \). The amount of variability in factor placement explained by group membership was 22.7%. While the OLD definition group had a negative placement on the POTENCY factor, the NEW definition was situated as significantly different and highly positive. The NEW definition was seen as highly substantive or potent relative to the OLD definition; that is, it is more exact, complete, objective, inflexible, safe, measurable and/or tangible.

The difference in placements on factor 2, EVALUATIVE, between the NEW and OLD groups, was found to be only weakly significant \( (F = 3.26, P = 0.072) \). The NEW definition was seen as somewhat ‘beneficial’ than the OLD definition.

On factor 3, ACTIVITY, the difference in placements was highly significant \( (F = 40.15, P < 0.001) \). For this factor, the NEW definition was seen as more negative. Relative to the OLD definition, it was seen as not accepting of activity. Given the difference between the OLD and NEW definitions can be viewed in terms of ACTIVITY (the recurring/non-recurring criterion), the negative relative placement on the ACTIVITY dimension for the NEW definition is logical. The NEW definition prohibits an EI from being of a recurring (i.e. active) nature.

It may be concluded that the measured meanings of the OLD and NEW definitions are significantly different. The groups had significantly different placements of the concept of EI across all three factors, although the important differences were on the POTENCY and ACTIVITY dimensions. The NEW definition is perceived as being much more potent, slightly better, and much less tolerant of activity than the OLD definition. It is therefore possible to conclusively reject the null hypothesis \( H_01 \). There is a significant difference (in fact, more than one significant difference) between the measured meaning of the concept EI depicted in the OLD definition and the NEW.

### 4.2. Analysis of classification decisions

This part of the data analysis deals with the analysis of the extraordinary/operating classification decisions made by subjects and tests whether there are significant between group differences in decision outcomes. The distribution of the responses over the treatment groups for the 10 cases is given in Table 4.

It can be seen that a more conservative decision pattern was exhibited by the group operating under the new definition; prima facie evidence that the experimental manipulation was effective. This pattern of responses is represented graphically in Fig. 2. Significant differences may be observed across the two subject groups for Cases 1, 2, 5 and 10. These cases all deal with recurring loss situations, the treatment of which was the focus of the change in definition.
The testing of H2 requires a comparison between the classification decisions prompted by the OLD and NEW Australian definitions. The analysis of variance, shown in Table 5, analyses the EI classification decision by the variable old/new definition for each of the ten cases. Consistent with Fig. 2, the cases for which significant differences were observed (Cases 1, 2, 5 and 10) were the examples of recurring losses; the type of cases where decisions made under the old definition were criticised as opportunistic. It seems that the new definition is “tougher” (more substantive, constraining or “potent”) on decisions where recurring items are involved.

Table 6 reports the results of an overall analysis of variance. The difference in decisions between the two groups of subjects prompted by the OLD and NEW definitions was significant ($F = 10.005$, $P = 0.002$). This evidence enables $H_{02}$ to be rejected. There is a significant difference between the EI classification decisions made by subjects on the basis of the OLD and NEW Australian definitions.

The conclusion, therefore, is that a significant change in decisions is apparent with the change in definition, but this change is, in part at least, a function of the individual differences in facts between the cases. Important components of the case facts are whether the item is of a recurring or non-recurring nature (the focus of the change in definition); and whether the item is a gain or loss (the focus of allegations of the differential treatment under the old definition). The $2 \times 2 \times 2$ ANOVA in Table 6 tests the impact of these variables. Whether an item is recurring or not is highly significant overall ($F = 51.643$, $P < 0.001$) in explaining the auditors’ EI decisions. Overall, there is no difference between the OLD and NEW definition groups in the classification of the non-recurring cases, while the NEW definition results in a significant increase in the number of recurring cases classified as operating.

### Table 4
Average case decisions by treatment group

<table>
<thead>
<tr>
<th>Case</th>
<th>Group 1 NEW</th>
<th>Group 2 OLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. By case (Requirement 2, Scale of 1 = Clearly extraordinary to 6 = Clearly operating)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Recurring/loss</td>
<td>4.30</td>
<td>3.45</td>
</tr>
<tr>
<td>2. Recurring/loss</td>
<td>4.90</td>
<td>3.70</td>
</tr>
<tr>
<td>3. Non-recurring/gain</td>
<td>3.05</td>
<td>3.15</td>
</tr>
<tr>
<td>4. Non-recurring/loss</td>
<td>3.95</td>
<td>3.90</td>
</tr>
<tr>
<td>5. Recurring/loss</td>
<td>5.55</td>
<td>4.40</td>
</tr>
<tr>
<td>6. Recurring/gain</td>
<td>5.40</td>
<td>5.20</td>
</tr>
<tr>
<td>7. Non-recurring/gain</td>
<td>3.25</td>
<td>3.37</td>
</tr>
<tr>
<td>8. Recurring/gain</td>
<td>4.65</td>
<td>4.55</td>
</tr>
<tr>
<td>9. Recurring/gain</td>
<td>4.55</td>
<td>4.65</td>
</tr>
<tr>
<td>10. Recurring/loss</td>
<td>4.65</td>
<td>3.45</td>
</tr>
<tr>
<td>B. Overall responses (Requirement 1: extraordinary/operating classification)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Extraordinary</td>
<td>12.5%</td>
<td>16.5%</td>
</tr>
<tr>
<td>% Operating</td>
<td>87.5%</td>
<td>83.5%</td>
</tr>
</tbody>
</table>

### Table 5
Analysis of variance: extraordinary/operating decision by variable old/new definition for each case

<table>
<thead>
<tr>
<th>Case number</th>
<th>Case facts</th>
<th>$F$</th>
<th>Sign of $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Recurring/Loss</td>
<td>3.15</td>
<td>0.084**</td>
</tr>
<tr>
<td>2</td>
<td>Recurring/Loss</td>
<td>7.6</td>
<td>0.009***</td>
</tr>
<tr>
<td>3</td>
<td>Non-recurring/Gain</td>
<td>0.048</td>
<td>0.828</td>
</tr>
<tr>
<td>4</td>
<td>Non-recurring/Loss</td>
<td>0.008</td>
<td>0.929</td>
</tr>
<tr>
<td>5</td>
<td>Recurring/Loss</td>
<td>8.411</td>
<td>0.006***</td>
</tr>
<tr>
<td>6</td>
<td>Recurring/Gain</td>
<td>0.543</td>
<td>0.466</td>
</tr>
<tr>
<td>7</td>
<td>Non-recurring/Gain</td>
<td>0.056</td>
<td>0.814</td>
</tr>
<tr>
<td>8</td>
<td>Recurring/Gain</td>
<td>0.062</td>
<td>0.805</td>
</tr>
<tr>
<td>9</td>
<td>Recurring/Gain</td>
<td>0.066</td>
<td>0.799</td>
</tr>
<tr>
<td>10</td>
<td>Recurring/Loss</td>
<td>7.653</td>
<td>0.009***</td>
</tr>
</tbody>
</table>

*Significant at 0.10 level.
**Significant at 0.05 level.
***Significant at 0.01 level.
items. Whether an item is a gain or loss is not significant as a main effect \((F = 1.606, p = .206)\). However, a significant interaction \((F = 7.538, P = 0.006)\) exists between the variables OLD/NEW definition and GAIN/LOSS. The nature of the interaction relates to the fact that the definition used had no impact on the decisions made for the gain cases; the difference measured was only in respect of losses. A number of cases designated as ‘extraordinary’ under the old definition are regarded as ‘operating’ items under the new definition. It is notable that EI classification decisions made under the NEW definition thus differ from those under the OLD definition where the case relates to a recurring item involving a loss. Other cases (GAIN/RECURRING, GAIN/NON-RECURRING, LOSS/NON-RECURRING) result in similar decisions between the two subject groups.

4.3. Association between measured meaning and decision outcomes

This part of the data analysis tests the proposition made in H3; that a significant amount of the variability in extraordinary classification decisions is associated with the variability in the meaning of the concept ‘extraordinary’ upon which those decisions are, in part, based. The decision outcomes are correlated with the factor placements (location of the concept within the cube). The results of the Spearman correlations are shown in Table 7.

These results provide evidence for the acceptance of H3. For members of both the OLD and NEW definition groups, the variability in decisions can be, in part, explained by the variability in measured meaning of the decision rule used. However, the level of explained variability is not high.\(^{17}\)

It can be observed from Table 7 that the relationship between measured meaning and decisions is driven by factor 1, POTENCY. The relationships of factor 2, EVALUATIVE, and factor 3, ACTIVITY, with classification decisions were not found to be significant. This is consistent with the results obtained in the measurement of meaning factor analysis. Further evidence is provided by

\(^{17}\) The factor scores vary by group. The semantic differential instrument was administered by group, rather than by decision, i.e. after all 10 cases and not after each case. A requirement to complete a semantic differential after each case would have resulted in subject resistance, boredom and fatigue. The relationship between factor scores and individual decisions is therefore a function of how much on average the total decisions changed because of the different decision rule. Therefore the variability in meaning in the rule will explain a relatively low proportion of the variability of individual case decisions. Where the variability in the meaning of the rule and the case facts are combined (as in the ordinal probit analysis shown in Table 8), a very high proportion of the variance is explained.
the results of an ordinal probit analysis, shown in Table 8.

Ordinal probit is the appropriate form of regression analysis, as the dependent variable (the classification decision) is categorical and scaled 1 to 6 from ‘Clearly Extraordinary’ to ‘Clearly Operating’. The significance of the threshold parameters validates the use of the technique. As expected, and consistent with the results in Table 7, the coefficient on the first factor score, POTENCY, is significantly negative; while the other two factors were not significant. The coefficient of the case facts variable (the categorical variable Case 1 to 10) is also highly significant. The observed relationship between measured meaning and decision outcomes is driven by the potency of the new definition and by the nature of the case facts in question.

### 5. Discussion

Based on the results reported here, three major conclusions may be drawn. These conclusions relate to (1) cognitive structures and meaning; (2) the classification decisions; and (3) the relationship between measured meaning and decision outcomes.

In relation to cognitive structures and meaning, the factor structure found to exist was consistent with the classic Evaluative–Potency–Activity structure that was originally proposed by Osgood et al. (1957) as the universal factor structure applicable to general judgments or the general domain of meaning; and demonstrated by Houghton (1988) to be the structure that underlies the specialised area of judgments that is the accounting domain of meaning. This three factor structure was established as being stable and robust. The conclusion that all subjects shared the same structure is important. It indicates that they all have a complex, multi-dimensional cognitive structure in respect of accounting concepts. This shared structure is not unexpected, given the homogeneous nature of the subjects (all practicing auditors). It is also important as it means that within the context of this shared cognitive structure, valid comparisons may be made between the
subjects’ placements of concepts (i.e. their perceived meanings).

Significant between-group differences in the measured meaning of the concept of EI were found to be present, leading to the rejection of Ho1. A highly significant difference in placement (i.e. measured meaning) was evident on the POTENCY dimension. The OLD definition was seen as more negative than the NEW definition. The NEW definition was therefore seen as substantially more potent, stronger or substantive than the OLD definition. On the EVALUATIVE dimension, the NEW definition was seen as only slightly better or more positive than the old definition. On the ACTIVITY dimension, the NEW definition was more negative than the OLD definition: the new definition was perceived as “hostile” to activity. As the difference between the OLD and NEW definitions can be viewed in terms of activity (the recurring/non-recurring criterion), the negative position of activity appears logical.

The OLD and NEW definitions of EI are therefore perceived as having significantly different meanings. The differences are perceived to lie in the increased potency or substantiveness of the new definition and its lack of tolerance to activity. Such a result is consistent with the nature of the changes made to the wording of the definition.

Turning to the classification decisions, significant between-group differences in decision outcomes were observed. Under the new definition, many more cases were designated ‘operating’ whereas under the previous definition they were designated ‘extraordinary’. The new definition induced a more ‘conservative’ decision pattern in subjects. A significant amount of variability in decision outcomes can be attributed to group membership, that is, OLD or NEW definition.

The major conclusion of the study deals with the relationship between the meaning of the definition used in the accounting standard and the classification decisions. Accounting judgments and decisions are influenced by many factors. In relation to the EI classification decision, some of those variables would include institutional factors such as the training programs and standard policies adopted by employer accounting firms; the incentives to which auditors and client management are subject; and facts about the particular event or transaction in question, including such things as the relative size of the item, the industry of the client, whether the item was a gain or loss and whether it was recurring in nature. Amongst the factors influencing the decision outcome is the meaning of the defined accounting concept upon which the decision is based.

We conclude that a change in the regulated definition is perceived as having a different meaning by one important party in the financial reporting process (auditors) and that there is a systematic relationship between the perceived meaning and the subsequent classification decision outcome, in the direction intended by the regulators.

6. Summary and conclusion

The nature of the accounting judgment examined means that the findings have implications for accounting policy making and practice. It is clear that changes in the regulatory environment do have a significant influence upon the judgments of accounting practitioners. Subtle changes in the wording of regulations are perceived to have differences in meaning, and these differences are detected and acted upon by practitioners.

Although this study was applied to a retrospective change in definition, the methodology is suitable to evaluate the potential impact on disclosure decisions of proposed accounting standards, and to choose amongst alternative wordings of proposed standards or regulations. This can assist in accounting standards operating more effectively as a constraint on aggressive reporting, in tandem with the role of the auditor as a monitor of the financial statements. A sub-set of the earnings management and audit judgment literatures has dealt with practitioners’ interpretations of probability phrases used in accounting and auditing standards (for example, Amer, Hackenbrack & Nelson, 1994; Reimers, 1992). Recent studies (Cuccia, Hackenbrack & Nelson, 1995; Hackenbrack & Nelson, 1996) have found that auditors justified “aggressive” as well as “conservative” reporting decisions on the basis of their interpretation of vague probability phrases in

financial accounting standards, when driven by economic incentives to do so. The flip-side of this is the proposition that carefully-worded standards and definitions can assist auditors in justifying judgments that are in keeping with the spirit of the regulation. In this study, the wording of the regulated definition of an extraordinary item was changed to become arguably more precise, with the intention of reducing flexibility in interpretation and hence “aggressive” reporting decisions. This study provides evidence that changing the wording of regulatory requirements may in fact mitigate “aggressive” reporting.

The findings of this study also have implications for the measurement of meaning in accounting literature. Previous studies in this literature have tended to focus on the meanings held by different subject groups (producers and users of accounting information) of various accounting concepts. The present study extends the previous work to examine the effect that different perceived meanings may have upon the judgments of individuals. The study of meaning in an applied decision context, and the establishment of a link between connotative meaning and decision outcomes, is a first in accounting research; and will add to the understanding of the processing of accounting information.

The findings must be viewed in the light of a number of limitations, some of which apply to experimental research generally; others apply to this study alone. The most important limitation is that it was only the connotative meaning of the applicable decision rule that was measured, rather than the meaning of the complete set of accounting information, or other aspects of meaning. The dichotomous classification of meaning as denotative/connotative has been criticised (see, for example, Garza-Cuaron, 1991; Harris, 1981). A potential shortcoming is thus the assumption that literal or denotative meanings are shared, and that the important variations in constructed meaning takes place in the connotative. However, this is likely to be mitigated by the nature of the subjects and of the study. The production of meaning is a continuous and dynamic process; meaning is culturally and historically located, contextually-embedded, and may change over time. As this study takes a cross-sectional rather than a longitudinal approach, variation in denotative or literal meaning is effectively controlled for. Admittedly, although we do not actually measure denotative meaning, it is a reasonable assumption that, for a group of audit professionals, all using the same technical professional definition, there is not likely to be any significant variation in literal meaning. Related to this, the semantic differential technique of Osgood et al. (1957) is based on the underlying assumption of a process model of communication. As a multi-disciplinary area of study, communication has been defined and modeled from various theoretical viewpoints; of which the Osgood model is but one. However, it does provide a means by which evidence may be obtained as to the proposition or belief that their perceived meanings influence subjects’ actions or decisions. Hence, the focus of the study is on the subjects (auditors); the message (the accounting standard definition); and the effect (the decisions made on the basis of that definition); that is, on the interpretation or process of constructing meaning.

The use of an experimental setting potentially restricts external validity, and the standard limitations of experimental research are acknowledged. In common with most experimental studies, convenience rather than truly random sampling was used, due to constraints on subject availability. The nature of the sample may also place some limitations on the generalizability of results; for example, they may not be generalizable to audit partners, or accountants outside of the Big 6 accounting firms, or to outside of the Australian regulatory environment.

While the cases used in this study are based on real examples and are representative of the types of decisions encountered in practice, they cannot represent all possible cases, nor the proportions they appear in practice, and were of necessity somewhat simplified. There may possibly have been an interaction between subjects’ perceptions of materiality levels and their meanings and/or decision outcomes. Also, cases were presented in a fixed order, which may possibly induce some kind of order bias.

Finally, there are many factors which contribute to the variability in decision outcomes, such as the decision context, institutional factors and indivi-
dual personality factors. This study focused on only one aspect of the communication of accounting information: that of its connotative meaning.

There are several potential avenues for future research. The meaning of other parts of the set of accounting information (e.g. the case facts) and their impact on decision outcomes could be measured. In the accounting context, there are many other types of decisions in which the role of connotative meaning could be important and could be tested. This includes, but is not limited to, accounting concepts defined in existing and prospective accounting standards. Only a single subject group, experienced practicing auditors, was used. The meaning of an EI may differ for different types of subjects, as these groups would have different decision contexts and be subject to different incentives. For example, the meaning of an EI could be perceived differently by a shareholder making an investment decision, than by a banker making a lending decision. One aspect giving rise to the construction of differing meanings is likely to be the economic incentives to which the individuals’ decisions are subject.

To conclude, the findings of this study, that changes in defined accounting rules significantly affect auditor judgments, is evidence of the propositions that meaning can influence judgments, and that careful playing of the “language game” by standard-setters has the potential to constrain aggressive financial reporting.

Acknowledgements

The authors acknowledge the financial assistance of the then Coopers & Lybrand Accounting Education and Research Fund, and thank the following firms for their co-operation: KPMG Peat Marwick, Arthur Andersen & Co, and the then Coopers & Lybrand and Price Waterhouse. The assistance of Kim Sawyer (University of Melbourne), and the comments of Nancy Bagranoff (American University), Paul Munter (University of Miami), Peter Luckett (University of New South Wales), Rashad Abdel-khalik (University of Florida), Mary Ellen Oliverio (Pace University), Len O’Neill (University of Melbourne), and participants at workshops at Curtin University of Technology, Griffith University and the European Accounting Association, Accounting Association of Australia and New Zealand, Accounting, Behavior and Organizations and American Accounting Association Conferences are also gratefully acknowledged, as are the comments of two anonymous reviewers.
Appendix. Extract from sample instrument

The following definition of extraordinary items was applied by you in the judgment tasks:

“Extraordinary items” means items of revenue and expense which are attributable to events or transactions of a type that are outside the ordinary operations of the company or group of companies and are not of a recurring nature.

In the context of the above definition, the term “EXTRAORDINARY ITEM” tends to be:

| Exact   | Estimated | Bad      | Good    | Measurable | Unmeasurable | Necessary | Unnecessary | Planned   | Unplanned | Objective | Subjective | Tangible  | Intangible | Strong  | Weak    | Indirect | Direct | Variable | Constant | Safe    | Risky   | Complete | Incomplete | Discretionary | Required | Imaginary | Real      | Adverse | Temporary | Permanent | Controllable | Unexpected | Expected | Active  | Dynamic | Short-term | Inflexible | Flexible |
|---------|-----------|----------|---------|-----------|-------------|-----------|------------|----------|-----------|-----------|------------|-----------|-----------|---------|---------|----------|--------|----------|-----------|--------|---------|----------|---------|-----------|-----------|------------|----------|----------|----------|---------|-----------|-----------|-------------|----------|---------|---------|---------|-----------|-----------|---

References


management negotiation about the client’s financial reporting. Working paper, University of Alberta.


