

The U.S. versus international ethics codes for accountants: A computerized content analysis

Hoseoup Lee. College of Business Management, SUNY Polytechnic Institute, USA, leeh@sunypoly.edu
Patrick Wing Yin Leung. Caritas Institute of Higher Education, HONG KONG, pleung@cihe.edu.hk
Gim S. Seow. Milgard School of Business, University of Washington at Tacoma, USA gseow@uw.edu
Kinsun Tam. University at Albany, SUNY, USA, tam@albany.edu

Abstract. This study applies computerized content analysis techniques, including term frequency analysis, readability index, and cosine similarity, to compare the U.S. versus international ethics codes for professional accountants. It reveals and compares new characteristics of these ethics codes. The U.S. ethics code is found to be less readable but more reflective of moral values concerning social order. The international ethics code emphasizes the independence and confidentiality principles, while the U.S. ethics code emphasizes independence and responsibilities. There are more similarities within the U.S. ethics code and within the international ethics code than between these ethics codes.

Keywords: ethics codes for accountants, computerized content analysis, term frequency, readability, cosine similarity

Acknowledgement: The authors are grateful for constructive comments from K. Choong, Michael P. Coyne, Joe Z. Shangguan, and two anonymous reviewers on an early version of this paper.

1. INTRODUCTION

International Code of Ethics for Professional Accountants (hereinafter international ethics code) represents the base case ethical requirement in any IFAC (International Federation of Accountants, 2018) member country. The national ethics code is one adopted by a given nation's professional accounting body for compliance by its members. These two ethics codes are not necessarily identical. Accounting firms serving multinational clients need to be aware of ethical expectations within their operational and regulatory environment, which vary depending on the jurisdiction governing the firms' work.

A national ethics code for accountants may deviate from the international ethics code if (1) the IFAC member organization adopts the international ethics code with modifications, or (2) the IFAC member organization chooses not to adopt the international ethics code, but instead creates its own, more stringent, ethics code (Fatemi et al. 2020). In the U.S., the AICPA (American Institute of Certified Public Accountants) drafts and adopts its own ethics code (i.e. AICPA Code of Professional Conduct 2016, hereinafter U.S. ethics code). While the U.S. ethics code details the ethical expectations on U.S. accountants, such expectations, buried in 187 pages of text, are inevitably less than concise. We argue that a comparison of ethics codes can yield new insights. One quick way to highlight the unique characteristics of the U.S. ethics code is by comparing it to another ethics code. The international ethics code is a meaningful and convenient benchmark for this purpose.

The coverage of the academic literature on ethics codes for accountants is sparse and unbalanced. In particular, comparison of ethics codes in accounting is a largely unexplored research topic, let alone comparing characteristics of national versus international ethics codes by means of computerized content analysis. This study applies computational linguistics techniques, in particular, term frequency, readability, and cosine similarity (defined later in this section), to U.S. and international ethics codes to help address these research questions: (1) How different are the emphases of these two ethics codes? (2) How readable are these ethics codes, and how different are their readability? (3) To what extent are these ethics codes similar to each other, and how are their sections and subsections related?

We find that accountants in public practice, whether in the U.S. or internationally, are subject to more lengthy ethics code than accountants in business practice. When compared to the international ethics code, the U.S. ethics code is less readable and more focused on moral values concerning social order. The U.S. ethics code stresses independence and the responsibilities of the accountant, while the international ethics code focuses on the independence and confidentiality principles. Moreover, the emphasis of the U.S. ethics code on objectivity and independence considerably exceeds that of the international ethics code. Finally, we note more similarities within the U.S. and the international ethics codes than between these ethics codes.

This study complements prior research with its unique focus and innovative methodology. Our results reveal similarities and differences between U.S. and international ethics codes not identified in prior research. Accountants must understand the ethical requirements in various jurisdictions in which they operate and report. Investors comparing investment opportunities across countries must assess the quality of financial statements with respect to the ethical codes guiding the accountant's and the auditor's work on such statements. Furthermore, comparing accounting ethics codes using computerized content analysis tools introduces a new research direction. Procedures highlighted in this study to prepare ethics code text data for analysis are useful for future research.

We organize the rest of this article in the following manner. Section 2 reviews the literature on professional ethics codes in accounting. Section 3 describes the research methodology and text analytic measures. Section 4 explains research data preparation and defines variables. Section 5 reports the results of analyses. Section 6 discusses the results and relates them to the literature. Conclusions, contributions, limitations, and future research are presented in Section 7.

2. LITERATURE REVIEW

This section reviews the literature on accounting ethics code and textual analysis tools used in this study.

2.1. Review of Accounting Ethics Code Literature

The current literature on accounting ethics codes revolves around three themes, namely ethics education, usefulness of ethics codes, and qualitative comparisons

between national and international ethics codes. The following subsections describe the extent of prior studies conducted in these three areas.

2.1.1. Ethics education

One major stream of accounting ethics research focuses on ethics education (Cooper et al. 2008). Ethical judgement on accounting issues is relevant not only to professional accountants in practice, but also to the training of accounting students on ethical awareness (Stanga & Turpen, 1991). According to Shawver and Miller (2017), ethics content in an accounting class positively influences the students' perception of moral intensity.

The mainstream accounting curriculum is lacking non-Western ethical theories and contemporary Western ethical thought (Tweedie et al., 2013). Cheung and Agrawal (2018) question the level of ethics emphasis in accounting curriculum and in the recruitment of professional accountants. Insufficient ethics training and ineffective ethics education in accounting are blamed for the prevalence of ethical violations and frauds in the accounting profession (Pergola & Walters, 2017; Miller & Shawver, 2018). Despite formal ethics education and ethics-related professional training, professional accountants continue to play a central role in enabling corporations to make unethical business decisions (Reinstein et al., 2019).

An extensive body of literature suggests various means for improving ethics training in accounting (Pergola & Walters, 2017). Fischer and Rosenzweig (1995) suggest increasing student and practitioner sensitivities to ethical issues. Mele (2005) encourages acquisition of moral values and virtues in ethical education in accounting. Christensen et al. (2018) suggest incorporating action-oriented instruction in accounting ethics education. The accountant's ethical training should focus on ethical reasoning rather than memorizing ethical standards (Reinstein et al., 2019). The IFAC's eCode initiative leverages web-based information technologies to train professional accountants in navigating and accessing key ethics and independence provisions (Pieters, 2019).

2.1.2. Usefulness of ethics codes

The accounting profession believes in monitoring the ethical behavior of its members through self-regulation and peer review (Cohen & Pant, 1991). Accounting ethics research examines the accountant's conformity with

professional ethics and recommends appropriate ethical conduct (Flory et al., 1992). The literature acknowledges the usefulness of professional ethics codes. Communication and reinforcement of an ethical code of conduct affect the quality of audit judgment (Martinov-Bennie & Pflugrath, 2009). The code of conduct in accounting is an important control against ethical violations by accountants (Cardona et al., 2019). Activated ethical norms can curb aggressiveness of accounting decisions (Kelly & Murphy, 2019).

However, there are deficiencies in ethics codes. Spalding and Oddo (2011) find the AICPA Code of Professional Conduct to be inadequate or poorly crafted with respect to five general principles of the IFAC's Code of Ethics, namely objectivity, integrity, inquisitiveness, loyalty, and trustworthiness. Ethics codes cannot prevent ethical failures. Large-scale financial scandals can be traced to ethical lapses, which result in the joint manager-accountant disregard of their fiduciary responsibilities to the client and the public (Curtis, 2008; Goel, 2019).

2.1.3. Similarities and differences

Allen (2010) finds the U.S. and international ethics codes to be more similar than different. Such similarities are the consequences of the composition of IFAC's membership. Member organizations of the IFAC, such as the AICPA and the IMA, must either adopt the IFAC Code of Ethics (perhaps with minor modifications) or adopt standards at least on par with the IFAC Code of Ethics (Clements et al., 2009a). So American CPAs and CMAs must somehow comply with international ethics standards (Allen, 2018; Anders, 2018).

Cultural factors are key drivers behind the differences between international and national ethics codes. Reflecting a shared core agreement, international standards must not impose the tradition of one culture on all others (Tweedie et al., 2013). However, local cultural or religious practices produce highly diverse and deeply held ethical beliefs (Tweedie et al., 2013) such that ethical perceptions and standards vary across cultures and countries (Cohen et al., 1993). In fact, cultural factors affect the acceptance and implementation of international code of conduct in accounting (Cohen et al., 1992). Clements et al. (2009b) find about 50% of IFAC member organizations having adopted the IFAC Code of Conduct, with cultural characteristics influencing the adoption decision.

Multinational accounting firms need to anticipate international differences in ethical behavior (Cohen et al., 1996) and manage a multitude of ethical perceptions and standards to ensure optimal audit quality and ethical conduct (Cohen et al., 1993). Cohen et al. (1993) review professional ethics in the international accounting profession, highlighting differences in ethical perceptions driven by cultural differences among countries. Applying Hofstede's (1980, 1991) cultural dimensions theory, Cohen et al. (1993) identify culture-based ethical conflicts in international accounting and auditing practices. The accounting ethics curriculum, Tweedie et al. (2013) argue, should accommodate variations in ethical traditions and practices between nations and cultures.

Another difference appears in tax accounting. Ethical expectation on the tax accountant based on the IFAC's professional integrity principle differs from the AICPA's client advocacy principle (Fatemi et al., 2020). According to the IFAC's NOCLAR (Non-compliance with laws and regulations) provisions, professional accountants, regardless of their capacity, must not turn a blind eye to illegalities committed by the client or employer. However, tax accountants in the U.S. are expected to be client advocates rather than professionals with independent judgement (AICPA, undated). NOCLAR provisions are not yet incorporated into the AICPA code of professional conduct (Anders, 2018). U.S. tax accountants may have difficulty balancing their evidence evaluation roles and their primary responsibility as client advocates (Bobek et al., 2010).

2.2. Textual analysis tools

Conducting computerized content analysis on ethics codes differentiates this study from the extant literature. Various textual analysis tools are used in accounting and business research, but rarely in the analysis of ethics codes. In particular, term frequency, readability, and cosine similarity are widely used to compare business documents.

2.2.1. Term frequency

Terms appearing at high frequencies naturally reflect the author's emphases. Term frequency analyses, which link a higher frequency of discussion to greater author attention, is commonly used in managerial attention studies (Sonpar & Golden-Biddle, 2008; Hu et al., 2018; Audi et al., 2016; Yang et al., 2017). Term frequency scaled by total text length is a measure of emphasis established in

Tetlock et al. (2008), Kothari et al. (2009), and Matsumoto et al. (2011). Tetlock et al. (2008) and Kothari et al. (2009) laud the term frequency measure as "parsimonious, objective, replicable, and transparent", and analysis based on term frequencies as "comprehensive" and "amenable to replication in future". Subsequent research (Mayew & Venkatachalam, 2012; Chen et al., 2014; Heston & Sinha, 2015) continues using the term frequency measure. By comparing high frequency terms, this study provides objective evidence on the similarities and differences of the ethics codes' emphases.

2.2.2. Readability

A major application of readability in accounting research is to address the SEC's (2007) concern that management conceals fraud beneath verbose reports and disclosures. Evaluating the link between financial report readability and profitability, Li (2008) finds financial reports of less profitable firms to be less readable. Extending this line of research, Lo et al. (2017) document a marginal link between poor readability and future financial misstatements, and find that firms inflating earnings to meet or beat analyst forecasts present less readable financial reports. On the other hand, Lundholm et al. (2014) argue that clearer, more readable, and more concrete corporate disclosures reduce the psychological distance between the firm and the investor, thus encouraging investment in the firm's stocks.

2.2.3. Cosine Similarity

Brown and Tucker (2011), using cosine similarity as a measure of modification, study the extent of year-over-year modifications of MD&As (management discussion and analysis). In Hanley and Hoberg (2012), cosine similarity measures the amount of new information between old and new IPO prospectuses. Ball et al. (2015) identify peer firms based on the cosine similarity measure computed over MD&As. Lang and Stice-Lawrence (2015) compare annual reports of a given firm and its foreign peers based on cosine similarity. In Jung (2018), cosine similarity computed between a firm's MD&A and a given finance (innovation) textbook is adopted as a proxy for the CEO's level of emphasis on finance (innovation).

2.3. A quick summary

The above review provides a snapshot of research work on accounting ethics and textual analysis tools used in accounting and business research. Extensive research on ethics codes focuses on education and training in accounting. Some research examines the usefulness and limitations of ethics codes. Research on the similarities and differences between international and national ethics codes focuses largely on the cultural and tax dimensions. We argue that many similarities and differences remain unclear because of the current literature's limited and highly selective focuses. Therefore, we continue this line of research with an open mind. By conducting computerized content analysis on the term frequencies, readability, and cosine similarity of U.S. and international ethics codes, which has not been attempted in prior studies, we aim at identifying new similarities and differences.

3. RESEARCH METHODOLOGY AND TEXT ANALYTIC MEASURE¹

In this study, multiple computer programs are written to compute term frequencies, readability indexes, and cosine similarity values of the U.S. versus international ethics codes. These measures are analyzed to illustrate similarities and differences between the two ethics codes.

Term frequency analysis characterizes a document based on the number of times (i.e. frequency) that each term appears in the document. To evaluate a specific emphasis on a given topic, term frequency analysis is refined to focus on a predetermined set of terms (i.e. a dictionary) related to the topic. To control for variations in number of terms driven by document lengths, term frequencies are often scaled by some measure of document length.

Readability is the ease with which a reader can understand a written text (Wikipedia). Readability is regarded as the first consideration for a code of ethics for the CPAs (Calhoun et al., 1998). Ideally, the ethics code should be easily readable and understandable to everyone (Calhoun et al., 1998). Therefore, comparing the readability of ethics codes is meaningful. Li (2008) uses both the Fog index (FOG) and document length of the financial report as measures of readability. Loughran and McDonald (2014) advocate using document length as a

¹ The review of relevant computerized content analysis literature in accounting and business is not meant to be exhaustive. For a detailed account of the relevant literature, the reader may want to refer to Loughran and McDonald (2016).

measure of the readability of 10-K reports. This study adopts both the Fog index and document length of ethics code as readability proxies.

FOG, the most widely used readability index for calibrating English text readability, is designed to proxy for the number of years of formal education required to understand a given document in first time reading. Hence, a lower FOG indicates a more readable text. FOG is numerically given by the following formula:

$$\text{FOG} = (\text{words_per_sentence} + \text{percent_complex_words}) \times 0.4$$

Moreover, Ball et al. (2015) use the dimension measure, defined as the number of distinct words in a document, to gauge the size of vocabulary. According to Dolch (1949) and Stahl (2003), a robust relationship exists between vocabulary size and difficulties in reading. As in Xu et al. (2018), this study adopts dimension as another measure of readability, where a higher dimension corresponds to more difficult reading.

Cosine similarity is a measure of the similarity between two documents represented in the vector space model.² When a document is represented as a vector, each keyword becomes a dimension and the value of the dimension is the number of occurrence of the corresponding keyword in the document.³ The smaller the angle between two vectors, the more similar the two underlying documents. The cosine value of this angle is bounded on the closed interval [0, 1]. When two documents share no common keywords, their two corresponding vectors are orthogonal, yielding a cosine similarity of cosine ($\pi/2$) or 0. Conversely, when two documents contain the same frequencies of keywords, the corresponding vectors point in exactly the same direction, shrinking the angle between them to zero. Cosine similarity, calculated as cosine (0), thus assumes the value of 1.⁴ Therefore, a higher cosine similarity value indicates greater similarity between two documents.

² See https://en.wikipedia.org/wiki/Cosine_similarity for an introduction to cosine similarity

³ An alternative calculation of cosine similarity is based on tfidf (term frequency * inverse document frequency), where the frequency of each keyword is weighted by the inverse document frequency. Document frequency is the number of documents containing a given keyword. Inverse document frequency is the reciprocal of document frequency.

⁴ Two vectors do not have to be equal for cosine similarity to be 1. As long as two vectors point in the same direction (i.e. two vectors being collinear), the angle between them is zero, yielding a cosine similarity of 1. To illustrate this property, suppose Document A says "Accountants should maintain objectivity." Suppose Document B says the same twice, "Accountants should maintain objectivity. Accountants should maintain objectivity." Although the vector representing B is twice as long as the vector representing A, they point in the same direction. The cosine similarity between the two documents is thus 1.

Manual content analysis, when applied to determine if two documents are similar or different, is short on quantifying the extent of similarity or difference. Cosine similarity addresses this limitation.⁵ Applying cosine similarity to the US versus international ethics codes enables us to quantify the similarity between corresponding sections of these ethics codes. Therefore, cosine similarity can direct us to sections of the US and international ethics codes with the highest similarity, or highlight sections with the biggest differences. International accounting firms, when serving clients in multiple jurisdictions, will find cosine similarity useful for identifying sections of ethics codes with high similarities or major differences.

Our computer programs are written in Java. Term frequency calculation is facilitated by Java's String class. The programs to compute readability and cosine similarity are respectively augmented by the Java version of the `Lingua::EN::Fathom` package and Apache Software Foundation's Lucene package.

To compute term frequencies related to a specific topic (such as moral values), the program searches the given document for each target term. The frequencies of all target terms are summed together. This total frequency measure, as an indication of overall emphasis on the topic, is then compared across documents.

To compute the FOG readability index, the program uses the `words_per_sentence()` method and `percent_complex_words()` method of the `Lingua::EN::Fathom` package. These methods count the number of words in each sentence and the number of complex words (i.e. those with three or more syllables) in each document. The outputs from these two methods are used to compute FOG according to the formula described earlier.

To compute cosine similarity, the two documents to be compared are first converted into two vectors. To compute a vector, the program identifies all unique words in a given document and counts the frequency of each word. We choose a version of cosine similarity that adjusts the frequency values with the tfidf (term frequency * inverse document frequency) weighing scheme as described in Manning and Schütze (1999) and advocated by Loughran and McDonald (2011). Apache's `dotProduct()` method computes cosine similarity as the dot product of the two vectors.

⁵ By definition, cosine difference (which is $1 - \text{cosine similarity}$) is a measure of difference between two documents.

4. RESEARCH DATA PREPARATION

This study computationally compares the U.S. ethics code for professional accountants (i.e. AICPA Code of Professional Conduct) versus international ethics code (i.e. International Code of Ethics for Professional Accountants). The texts of these two ethics codes are our source documents. We compare the entire documents as well as specific sections such as those to be applied to accountants in public practice, accountants in business practice, accountants in all practices, etc. Since the two ethics code are organized differently, sections bearing seemingly similar labels may contain different content. As explained in the subsequent paragraphs, some sections are reorganized to eliminate such variations for the sake of fair and meaningful comparison. Major sources of variation are summarize below.

1	The U.S. ethics code contains a specific section for members not in public practice or business, which is missing from the international ethics code.
2	The U.S. ethics code separates conceptual framework for members in business from conceptual framework for members in public practice. The international ethics code applies a single conceptual framework to both accountants in business and accountants in public practice.
3	The U.S. ethics code embeds the independence section within the section for members in public service, but the international ethics code maintains two top level sections on independence.
4	The international ethics code separates independence into two (for auditing and review engagements versus assurance engagements). The U.S. ethics code makes no such separation.
5	The U.S. ethics codes for publicly practicing members and business members each contains a section on ethical conflicts, but with essentially identical content. The international ethics code contains no separate section on ethical conflict.

We adopt a straightforward convention to label various sections of the ethics codes. The following table summarize these labels. The "us" and "in" prefixes are to identify the U.S. versus the international ethics codes. "Body" refers to the entire ethics code but with the cover page, publisher information, copyright, changes from previous edition, structure of presentation, table of contents, and appendixes removed. The "Body" of each ethics code contains multiple top level sections, which are highlighted in italics in the tables.

U.S. ethics code body and top level sections:	
<i>usBody</i>	The body of the U.S. ethics code.
<i>usCodeAll</i>	The section of the U.S. ethics code for all members.
<i>usCodeBusiness</i>	The section of the U.S. ethics code for members in business.
<i>usCodePublic</i>	The section of the U.S. ethics code for members in public practice.
<i>usCodeOther</i>	The section of the U.S. ethics code for other members.
U.S. ethics code subsections:	
<i>usConceptBusiness</i>	The U.S. conceptual framework for members in business.
<i>usConceptPublic</i>	The U.S. conceptual framework for members in public practice.
<i>usIndependence</i>	The subsection of the U.S. ethics code on independence.
<i>usCodePublicNet</i>	The subsection of the U.S. ethics code for members in public practice with the conceptual framework removed.
<i>usCodeBusinessNet</i>	The subsection of the U.S. ethics code for members in business with the conceptual framework removed.
International ethics code body and top level sections:	
<i>inBody</i>	The body of the international ethics code.
<i>inCodeAll</i>	The section of the international ethics code for all members.
<i>inCodeBusiness</i>	The section of the international ethics code for members in business.
<i>inCodePublic</i>	The section of the international ethics code for members in public practice.

<i>inIndependenceAudit</i>	The section of the international ethics code for audit and review engagements.
<i>inIndependenceAssurance</i>	The section of the international ethics code for assurance engagements.
International ethics code subsections:	
inConcept	The conceptual framework of the international ethics code.
inCodePublicCom	An artificial subsection of the international ethics code created from inCodePublic, inConcept, inIndependenceAudit, and inIndependenceAssurance for comparison purpose.
inCodeBusinessCom	An artificial subsection of the international ethics code created from inCodeBusiness and inConcept for comparison purpose.

The body of the U.S. ethics code (labelled in this study as usBody) contains top level sections including the section for all members (usCodeAll), the section for members in business (usCodeBusiness), the section for members in public practice (usCodePublic), and the section for other members (usCodeOther).

To enrich the scope of this study, we also analyze several subsections of the U.S. ethics code. For instance, the U.S. section on independence (usIndependence), contained in usCodePublic, is a subsection. In addition, the U.S. conceptual framework for members in business (usConceptBusiness) and the U.S. conceptual framework for members in public practice (usConceptPublic), embedded respectively in usCodeBusiness and usCodePublic, are subsections. For meaningful comparison, we remove usConceptBusiness and usConceptPublic from usCodeBusiness and usCodePublic to create usCodeBusinessNet and usCodePublicNet. To determine the extent that a specific conceptual framework guides the corresponding ethics code, we will compute the cosine similarity value between usConceptBusiness and usCodeBusinessNet (or between usConceptPublic and usCodePublicNet). Use of usCodeBusinessNet and usCodePublicNet reduces spurious similarity.

Similarly, the body of the international ethics code (inBody) contains top level sections (again highlighted in italics in the tables) including the section for all

members (`inCodeAll`), the section for members in business (`inCodeBusiness`), the section for members in public practice (`inCodePublic`), the section on independence for audit and review engagements (`inIndependenceAudit`), and the section on independence for assurance engagements (`inIndependenceAssurance`).

Likewise, we include the following subsections of the international ethics code to extend the scope of our analysis. The international ethics code's conceptual framework (`inConcept`), embedded in `inCodeAll`, is a subsection. As the U.S. ethics code for publicly practicing members (`usCodePublic`) contains also the conceptual framework and the section on independence, we create a compatible subsection for the international ethics code (i.e. `inCodePublicCom`, which sums `inCodePublic`, `inConcept`, `inIndependenceAudit`, and `inIndependenceAssurance`) for subsequent comparative analyses. Again, because the U.S. ethics code for business members (`usCodeBusiness`) contains the conceptual framework (but not the section on independence), we create a compatible section for the international ethics code (i.e. `inCodeBusinessCom`, which sums `inCodeBusiness` and `inConcept`) for subsequent comparative analyses.

5. RESULTS FROM COMPUTERIZED CONTENT ANALYSIS

5.1. Word frequency and readability

Table 1 presents the document length, word frequency, and readability of the ethics codes and their sections. The original size of the ethics codes and their sections are reported as `rawByteSize` in Column 1. Texts conveying supplementary information such as code section numbers and cross-referencing pointers are removed or "cleaned" to arrive at `cleanByteSize` (Column 2)⁶. Subsequent computation for word count, readability and cosine similarity is based on post-cleaning text. In Columns 1, 2, and 3, values of `usBody` and `inBody` approximate the sum of the top level sections. Discrepancies between `usBody` or `inBody` and their top level section sums are attributable to (a) tables of section page numbers, (b) glossary and abbreviation lists, and (c) references to nonauthoritative, new, revised, pending, and deleted guidance, which are in the "Body" but outside of the top level sections.

⁶ A code section number example is as in "Independence Rule [1.200.001]". A cross-referencing pointer example is "[Prior reference: paragraphs .134-.135 of ET section 191]".

	1	2	3	4	5	6	7	8	9
document	rawByteSize	cleanByteSize	cleanWordCount	cleanEthicCount	scaledEthicFreq	dimension	wordsPerSentence	percentComplexWords	fog
U.S.									
usBody	536288	469441	71975	854	0.0119	3417	28.71	28.61	22.93
<i>usCodeAll</i>	57389	49585	7648	112	0.0146	1424	20.70	27.73	19.37
<i>usCodeBusiness</i>	64010	57804	8662	152	0.0175	1182	29.13	30.66	23.91
<i>usCodePublic</i>	389895	346472	53315	496	0.0093	2976	31.38	28.47	23.94
<i>usCodeOther</i>	8229	7066	1088	27	0.0248	356	30.69	27.19	23.15
International									
inBody	457173	418378	63785	433	0.0068	2712	21.80	28.89	20.28
<i>inCodeAll</i>	31360	29290	4182	60	0.0143	776	20.36	33.63	21.60
<i>inCodeBusiness</i>	72465	67386	9841	91	0.0092	1298	19.78	32.28	20.83
<i>inCodePublic</i>	91502	85093	12845	76	0.0059	1363	21.07	29.89	20.38
<i>inIndependenceAudit</i>	158590	145483	22969	112	0.0049	1679	22.53	25.62	19.26
<i>inIndependenceAssurance</i>	70427	64800	10030	69	0.0069	1021	22.41	28.33	20.30

Table 1. Document Lengths, Ethics Word Frequency, and Readability of U.S. versus International Ethics Codes

Accountants in public practice and accountants in business are two major focuses of both the U.S. and the international ethics codes. In both sets of ethics code, the sections for accountants in public practice (i.e. *usCodePublic* and *inCodePublic*) are more voluminous than the corresponding sections for accountants in business (i.e. *usCodeBusiness* and *inCodeBusiness*) in terms of *rawByteSize*, *cleanByteSize*, and *cleanWordCount* (Columns 1, 2, and 3). Understandably, there are more expectations on ethical conduct of publicly practicing accountants who are answerable to the public.

In Column 4, *cleanEthicCount* refers to the count of ethics-related words. The *RcEthic* list of Harvard's *Inquirer* dictionary contains 151 word-senses relating to ethics, and in particular "moral values concerning social order".⁷ This list shrinks

⁷ See <http://www.wjh.harvard.edu/~inquirer/RcEthic.html>

to 138 ethics words after multiple senses of each word are condensed into a single word.⁸ Column 4 counts the occurrence of these words while Column 5 scales these counts by the overall word counts (i.e. by Column 3). In general, more ethics-related words appear in the U.S. ethics code than the international ethics code. The three top level sections containing the highest proportions of ethics words are all in the U.S. ethics code. These sections are (1) usCodeOther, (2) usCodeBusiness, and (3) usCodeAll.

It is noteworthy that scaledEthicFreq (Column 5) never reaches 3% in any row, reflecting the gap between general social conduct and the more specific accounting professional ethics. In particular, concepts important to the accountant's professional conduct, such as objectivity, undue influence, independence, threat, safeguard, confidentiality, etc., are rarely the emphases of a higher order social conduct.

Columns 2, 6, 7, 8, and 9 compare the readability of the ethics codes. Column 2 indicates longer U.S. ethics code than international ethics code, suggesting the U.S. ethics code being more unreadable per Loughran and McDonald (2014). Likewise, Column 6 reports generally higher dimension values, and hence higher reading difficulty, of U.S. ethics code than international ethics code. Columns 7-9 are based on the Fog Index. Columns 7 and 8 report the two components of FOG, namely words per sentence and percentage of complex words. The Fog index (Column 9) generally identifies the U.S. ethics code and top level sections as more difficult to read than the international ethics code and top level sections, an observation consistent with interpretations based on the document length and dimension. Among usBody, inBody, and all top level sections of the two set of ethics codes, the three most unreadable ones are all in the U.S. Code. They are (1)

⁸ For instance, two occurrences of "just" corresponding to its adjective and adverb forms are collapsed into one. The final list of Harvard Inquirer ethics words includes: "adhere", "adherence", "appall", "betray", "betrayal", "better", "board", "bound", "brave", "bravery", "broke", "cause", "censorship", "character", "cheat", "clean", "commitment", "confess", "confession", "conscience", "corrupt", "corruption", "crime", "criminal", "decency", "delinquency", "delinquent", "deplorable", "deplore", "desert", "deserve", "ding", "dirty", "duty", "earnest", "equitable", "ethical", "ethics", "fair", "faith", "false", "fine", "forgiven", "fraud", "goodness", "goodwill", "greed", "guarantor", "guilt", "guilty", "have", "heroism", "honest", "impartial", "impostor", "incumbent", "indictment", "indignant", "inequitable", "inequity", "infamous", "injustice", "innocence", "innocent", "integrity", "irresponsible", "just", "justice", "justifiable", "justifiably", "justification", "liar", "liberate", "lie", "manner", "manslaughter", "misbehavior", "mistreatment", "moral", "morality", "murderer", "notorious", "oath", "obligation", "oblige", "offence", "outlaw", "outrage", "penalty", "pledge", "pretence", "pretension", "principle", "proctor", "promise", "rectitude", "regret", "responsibility", "responsible", "right", "righteous", "righteousness", "rightful", "rob", "rue", "satisfactorily", "scandal", "scandalous", "shrank", "shrink", "shrunk", "sincere", "sincerity", "sorry", "steal", "stole", "stolen", "suppose", "telling", "temptation", "theft", "thief", "thieves", "treason", "trust", "truthful", "unfair", "ungrateful", "unjust", "unjustified", "untrustworthy", "vice", "violate", "violation", "wicked", "worthiness", "worthy", and "wrong".

the ethics code for members in public practice (usCodePublic, FOG = 23.94), (2) the ethics code for members in business (usCodeBusiness, 23.91), and (3) the ethics code for other members (usCodeOther, 23.15). We next pool usBody (inBody) and the top level sections of the U.S. (international) ethics code together, and statistically compare the dimension and Fog index of these two pools in t-tests (2-tailed; assuming independent samples and equal variance). The difference in dimension is found to be insignificant (p-value = 0.527), but the difference in Fog is significant (p-value = 0.027), which identifies the U.S. ethics code as more difficult to read.⁹

5.2. Emphases on principles of professional conduct

Table 2 compares the emphases on various Principles of Professional Conduct both within and across jurisdictions. The U.S. ethics code includes seven principles. These are "Preamble", "Responsibilities", "Public Interest", "Integrity", "Objectivity and Independence", "Due Care", and "Scope and Nature of Services". The Preamble highlights the member's "obligation of self-discipline above and beyond the requirements of laws and regulations", and calls for "an unswerving commitment to honorable behavior, even at the sacrifice of personal advantage". The Scope and Nature of Services principle reminds members in public practice to "observe the Principles of the Code of Professional Conduct in determining the scope and nature of services to be provided". Neither the Preamble nor the Scope and Nature of Services introduces distinct attribute of conduct. Accordingly, we focus on the other five principles. To allow for separate references to "objectivity" versus "independence", we decouple "Objectivity and Independence" into two principles. As a result, the terms describing the Principles of Professional Conduct in the U.S. ethics code are "responsibilities", "public interest", "integrity", "objectivity", "independence", and "due care". Using regular expression, our computer program accommodates morphological variants such as "responsible", "responsibly", "responsibility", "public interests", "objective", "objectively", "independent", and "independently" in counting.

The five Fundamental Principles in the international ethics code are "Integrity", "Objectivity", "Professional Competence and Due Care", "Confidentiality", and "Professional Behavior". For comparison with the U.S. principles, we add

⁹ We acknowledge limitations of applying t-tests in this paper because of small samples and/or the inherent relationship between observations.

"Independence", separately described in Parts 4A and 4B of the international ethics code, as an additional principle. To allow for separate referencing, we decouple "Professional Competence and Due Care" into "Competence" and "Due Care". The terms describing Fundamental Principles in the international ethics code are thus "integrity", "objectivity", "competence", "due care", "confidentiality", "professional behavior", and "independence". Using regular expression, our computer program accommodates morphological variants such as "objective", "objectively", "competent", "competently", "confidential", "confidentially", "professional behaviors", "independent", "and "independently" in counting.

The top half of Table 2 presents the scaled frequencies of principles of professional conduct appearing in the U.S. ethics code. The usBody row evinces the overall emphases on individual principles. Of all the principles of conduct, the U.S. ethics code in general emphasizes independence the most (scaled frequency of 0.00781) and due care the least (0.00010). The top three emphases are on independence (in usCodePublic and usCodeAll) and responsibilities (in usCodeAll).

The bottom half of Table 2 presents the scaled frequencies of principles of professional conduct appearing in the international ethics code. The inBody row shows the overall emphases on individual principles. Among the principles of conduct, the international ethics code in general emphasizes independence the most (0.00419) and professional behavior the least (0.00020). The top three emphases are on independence (in inIndependenceAssurance and inIndependenceAudit) and confidentiality (in inCodeAll).

Notably, although both the U.S. ethics code and the international ethics code emphasize independence, the emphasis is stronger in the U.S. ethics code as reflected by the higher scaled frequency values. Moreover, the U.S. ethics code regards "objectivity" as a principle closely related to "independence", which explains why these two concepts are originally combined as a compound principle. If we recombine "objectivity" and "independence", their joint emphasis in the U.S. ethics code as reflected by scaled frequency will be even higher than that of the international ethics code.

To more systematically compare the emphases on principles of professional conduct between the U.S. and international ethics codes, we calculate t-tests on

the scaled-frequencies of principle-related terms in the entire documents (i.e. usBody versus inBody). Our first t-test (2-tailed; assuming independent samples and equal variance) compares the 6 scaled frequency values of the U.S. ethics code with the 7 scaled frequency values of the international ethics code. The difference between these two groups is not statistically significant (p-value = 0.330). In our second t-test (2-tailed; paired-sample), we compare only those principles common to both ethics codes (i.e. integrity, objectivity, independence, and due care). The difference between the U.S. and the international ethics codes on these four principles, on a pairwise basis, is also insignificant (p-value = 0.201). We therefore have no statistical evidence to suggest differences between the U.S. and the international ethics codes in the emphases on principles of professional conduct. Given the small sample size, the failure to observe statistical significance is not surprising.

document	integrity	objectivity	independence	due care	responsibilities	public interest	competence	confidentiality	professional behavior
U.S.									
usBody	0.00146	0.00164	0.00781	0.00010	0.00239	0.00022			
<i>usCodeAll</i>	0.00248	0.00262	0.00615	0.00092	0.00510	0.00105			
<i>usCodeBusiness</i>	0.00312	0.00323	0.00035	0.00000	0.00150	0.00012			
<i>usCodePublic</i>	0.00107	0.00128	0.00953	0.00000	0.00221	0.00013			
<i>usCodeOther</i>	0.00000	0.00000	0.00000	0.00000	0.00184	0.00000			
International									
inBody	0.00052	0.00071	0.00419	0.00030			0.00053	0.00077	0.00020
<i>inCodeAll</i>	0.00239	0.00239	0.00359	0.00143			0.00287	0.00383	0.00096
<i>inCodeBusiness</i>	0.00061	0.00030	0.00010	0.00061			0.00071	0.00102	0.00030
<i>inCodePublic</i>	0.00062	0.00062	0.00078	0.00047			0.00062	0.00171	0.00039
<i>inIndependenceAudit</i>	0.00013	0.00057	0.00592	0.00000			0.00026	0.00000	0.00000
<i>inIndependenceAssurance</i>	0.00030	0.00060	0.00927	0.00000			0.00000	0.00000	0.00000

Table 2. Scaled Frequencies of Principle-related Terms in U.S. versus International Ethics Codes

5.3. Cosine similarity

Table 3 reports cosine similarity values. Panel A compares sections of the U.S. ethics code. Cosine similarity values in red highlights indicate how closely the

U.S. ethics code's conceptual framework (i.e. `usConceptBusiness` and `usConceptPublic`) relates to other sections of the U.S. ethics code. Regarded as the guiding principle of the code of conduct, the conceptual framework is expected to exhibit high cosine values in relations to the corresponding code sections. Because cosine similarity is bounded on the closed interval $[0, 1]$, we use the midpoint (i.e. 0.5) as the threshold between high and low values.¹⁰ In fact, the cosine similarity values corresponding to document pairs (`usConceptBusiness`, `usBody`), (`usConceptPublic`, `usBody`), (`usConceptPublic`, `usCodeAll`), (`usConceptBusiness`, `usCodeBusiness`), (`usConceptPublic`, `usCodePublic`), (`usConceptPublic`, `usIndependence`), (`usConceptPublic`, `usCodePublicNet`), (`usConceptBusiness`, `usCodeBusinessNet`), and (`usConceptPublic`, `usCodeBusinessNet`) exceed the middle value of 0.50. These cosine similarity values support the general assumption that the U.S. ethics code's conceptual framework serves as the guiding principle in the development of the U.S. ethics code.

Notably, cosine values of (`usConceptBusiness`, `usCodeBusinessNet`) and (`usConceptPublic`, `usCodePublicNet`) are smaller than those of (`usConceptBusiness`, `usCodeBusiness`) and (`usConceptPublic`, `usCodePublic`). This observation is not surprising. As we explain earlier, we remove the respective conceptual framework section from `usCodeBusiness` and `usCodePublic` to create `usCodeBusinessNet` and `usCodePublicNet` to reduce spurious similarity. Accordingly, cosine similarity values based on "Net" sections are numerically smaller but conceptually more appealing.

Cosine similarity values in green highlights (Panel B) indicate how closely the international ethics code's conceptual framework (i.e. `inConcept`) relates to other sections of the international ethics code. Most of these values, such as (`inConcept`, `inBody`), (`inConcept`, `inCodeAll`), (`inConcept`, `inCodeBusiness`), (`inConcept`, `inCodePublic`), (`inConcept`, `inCodePublicCom`), and (`inConcept`, `inCodeBusinessCom`) exceed 0.50. Again, these values support the general assumption that the international ethics code's conceptual framework guides the development of the international ethics code.

¹⁰ While the threshold value of 0.50 admittedly has no theoretical justification, it is used in practice (Erkan 2004; Ramage et al. 2009; Gandy et al. 2017; RxNPL 2019).

In Panel C, the blue highlights track the cosine similarity values between comparable sections of the U.S. and international ethics codes. Only 2, namely (usBody, inBody) and (usCodePublic, inCodePublicCom), out of 10 document pairs produce cosine similarity values in excess of 0.50. The ratio of high cosine similarity values (i.e. those exceeding 0.5) to all cosine similarity values in Panels A, B, and C are 36/45, 27/36, and 9/90.¹¹ Collectively, the above observations suggest that there is a higher level of similarity within each of the U.S. ethics code and the international ethics code than between these ethics codes.

Notably, (usCodeBusiness, inCodeBusinessCom) and (usCodePublic, inCodePublicCom) produce higher cosine similarity values than (usCodeBusiness, inCodeBusiness) and (usCodePublic, inCodePublic). As explained earlier, the "Com" sections are created to make more compatible comparisons, which naturally give rise to higher cosine similarity values.

Panel A. Within U.S. ethics code

Tfidf Similarity	Cosine	usBody	usCodeAll	usCodeBusiness	usCodePublic	usCodeOther	usConceptBusiness	usConceptPublic	usIndependence	usCodePublicNet
usCodeAll	0.821									
usCodeBusiness	0.760	0.613								
usCodePublic	0.990	0.774	0.681							
usCodeOther	0.486	0.410	0.627	0.418						
usConceptBusiness	0.587	0.456	0.768	0.535	0.345					
usConceptPublic	0.726	0.589	0.611	0.714	0.297	0.742				
usIndependence	0.930	0.721	0.525	0.962	0.293	0.429	0.614			
usCodePublicNet	0.989	0.771	0.674	0.999	0.418	0.515	0.687	0.966		
usCodeBusinessNet	0.741	0.604	0.977	0.662	0.657	0.615	0.507	0.505	0.660	

¹¹An alternative threshold value of high versus low cosine similarity is 0.7071, which corresponds to cosine ($\pi/4$). As the angle between two vectors representing two documents is bounded on the closed interval $[0, \pi/2]$, $\pi/4$ is the angular mid value. Based on this new threshold, the ratio of high cosine similarity values (i.e. those exceeding 0.7071) to all cosine similarity values in Panels A, B, and C are 17/45, 18/36, and 0/90. Therefore, this alternative threshold value does not change our observation of a higher level of similarity within the U.S. ethics code and within the international ethics code than between these ethics codes.

Panel B. Within international ethics code

Tfidf Cosine Similarity	<i>inBody</i>	<i>inCodeAll</i>	<i>inCodeBusiness</i>	<i>inCodePublic</i>	<i>inIndependenceAudit</i>	<i>inIndependenceAssurance</i>	<i>inConcept</i>	<i>inCodePublicCom</i>
<i>inCodeAll</i>	0.652							
<i>inCodeBusiness</i>	0.689	0.782						
<i>inCodePublic</i>	0.848	0.761	0.815					
<i>inIndependenceAudit</i>	0.857	0.316	0.325	0.546				
<i>inIndependenceAssurance</i>	0.753	0.292	0.299	0.469	0.626			
<i>inConcept</i>	0.646	0.933	0.719	0.720	0.343	0.332		
<i>inCodePublicCom</i>	0.980	0.536	0.546	0.772	0.916	0.800	0.556	
<i>inCodeBusinessCom</i>	0.714	0.856	0.988	0.835	0.345	0.322	0.819	0.576

Panel C. Between U.S. and international ethics codes

Tfidf Cosine Similarity	<i>usBody</i>	<i>usCodeAll</i>	<i>usCodeBusiness</i>	<i>usCodePublic</i>	<i>usCodeOther</i>	<i>usConceptBusiness</i>	<i>usConceptPublic</i>	<i>usIndependence</i>	<i>usCodePublicNet</i>	<i>usCodeBusinessNet</i>
<i>inBody</i>	0.518	0.508	0.396	0.506	0.207	0.376	0.590	0.447	0.493	0.363
<i>inCodeAll</i>	0.310	0.322	0.360	0.279	0.223	0.320	0.351	0.229	0.271	0.337
<i>inCodeBusiness</i>	0.290	0.292	0.385	0.255	0.204	0.375	0.293	0.209	0.249	0.351
<i>inCodePublic</i>	0.403	0.388	0.329	0.391	0.189	0.282	0.473	0.332	0.380	0.311
<i>inIndependenceAudit</i>	0.485	0.469	0.306	0.487	0.139	0.299	0.560	0.438	0.475	0.278
<i>inIndependenceAssurance</i>	0.402	0.363	0.257	0.405	0.135	0.271	0.473	0.371	0.395	0.226
<i>inConcept</i>	0.321	0.295	0.359	0.298	0.162	0.376	0.412	0.260	0.287	0.318
<i>inCodePublicCom</i>	0.523	0.497	0.361	0.520	0.181	0.350	0.611	0.464	0.507	0.329
<i>inCodeBusinessCom</i>	0.312	0.308	0.399	0.277	0.205	0.394	0.334	0.231	0.270	0.361

Table 3. Cosine Similarity within and between U.S. and International Ethics Codes

6. DISCUSSION OF RESULTS

The preceding section finds more ethics-related words in the U.S. ethics code than the international ethics code, suggesting stronger emphasis on moral values of the U.S. ethical code. According to Fatemi et al. (2020), any IFAC member organization choosing not to adopt the international ethics code will create its own, more stringent, ethics code. Choosing to write its own ethics code, the AICPA understandably specifies more stringent ethics expectations. The U.S. ethics code's stronger emphasis on ethics words than the international ethics code is therefore not surprising.

We observe stronger emphasis of the U.S. ethics code than the international ethics code on the independence principle. Hofstede's (2001) and Hofstede et al. (2010) describes a theory of cultural values in six-dimension. The U.S. attains a score of 91, one of the world's highest, on the individualism dimension (Hofstede et al., 2010). A high individualism score represents the lack of interdependence among members of society, such that a person is to take care of only oneself. The natural inclination of the U.S. culture to be individualistic is consistent with the stronger emphasis of the U.S. ethics code on independence.

We observe very few instances of high cosine similarity values between comparable sections of the U.S. and international ethics codes. This result is new to the current literature. Based on manual content analysis, Allen (2010) finds the international and US ethics codes to be more similar than different. Our computerized content analysis, conducted on newer versions of ethics code than those in Allen (2010), documents only limited similarities between comparable sections of the U.S. and international ethics codes. Our observations, from both within and between ethics codes, update and extend Allen's (2010) finding.

7. CONCLUSIONS, CONTRIBUTIONS, LIMITATIONS, AND FUTURE RESEARCH

7.1. Conclusions

This research represents the first study to apply computerized content analysis to accounting ethics codes. It reveals previously undocumented characteristics of the U.S. and international ethics codes, and compares these characteristics. Our major findings include the following: (1) Accountants in public practice are subject to more lengthy ethics code than accountants in business practice, both in the U.S.

and internationally. The U.S. ethics code contains a higher percentage of ethics words as provided in Harvard's Inquirer dictionary, suggesting stronger emphasis on moral values concerning social order, than the international ethics code. Moreover, the U.S. ethics code is less readable than the international ethics code. (2) The U.S. ethics code emphasizes the independence and the responsibilities principles, while the international ethics code emphasizes the independence and the confidentiality principles. Emphasis of the U.S. ethics code on objectivity and independence considerably exceeds that of the international ethics code. (3) There are more similarities within the U.S. ethics code and within the international ethics code than between these ethics codes.

7.2. Contributions

This study informs accounting professionals, academics, investors, standard setters, and other stakeholders of the differences between the U.S. and the international ethics codes for accountants. In this study, computational linguistics techniques reveal stronger emphases of the U.S. ethics code than international ethics code on moral values concerning social order, and on objectivity and independence. These observations are useful to professional accountants working in an international setting, and to academics, investors, and other stakeholders who may have assumed otherwise.

Reinstein et al. (2019) advocate focusing on ethical reasoning rather than memorizing ethical standards in training accountants. The emphases of the U.S. versus international ethics codes highlighted by our computational linguistics-based methodology arguably represent underlying ethical reasoning. Therefore, this study is relevant to the accountant's ethical training from Reinstein et al.'s (2019) perspective.

Prior studies point to cultural differences as potentially causing variations in ethics codes across jurisdictions, but provide no quantitative measure of such variations. The notion of differences between ethics codes may not take root in the minds of professional accountants, academics, and investors when the extent of differences is not objectively measurable. Using computational linguistics, this study quantifies the similarities and differences between two major ethics codes, thus enabling more objective comparisons.

Readability of ethics code is relevant and informative to the standard setter. We report Fog index values of ethics codes between 19.26 and 23.94. According to Ertugrul et al. (2017, p. 817), a Fog index of 20, which implies requiring 20 years of formal education or roughly a PhD degree to understand a given text, is excessive. Standard setters may therefore want to consider simplifying ethics codes. Quantitative readability measures and targets help standard setters identify the sections where readability improvement is needed.

On the methodological front, this study expands the set of analytical techniques to be applied to studies of ethics code. This study concisely demonstrates the use of a multitude of computerized content analysis techniques, including keyword count, readability, and cosine similarity, in text-based research in accounting. In addition, this study highlights the importance of cleaning raw data before processing. Reorganizing sections and cleaning off code number, cross references, etc. of ethics codes to improve comparability are useful considerations for future researchers using similar text data. After all, 80% of the work in data analytics is typically about finding, cleaning, and organizing data (Bowne-Anderson, 2018; Boskou et al., 2019).

7.3. Limitations and future research

This paper is organized according to computerized content analysis techniques, from simple to complex ones. We start with document length and word count measures, and end with cosine similarity. This organization is chosen for the benefit of the reader, who may be interested in ethics codes but not familiar with computational linguistics. Because each technique is applied to a unique aspect of the ethics codes, our results may not seem smoothly coordinated.

Computerized content analysis is not flawless. Keyword frequency, FOG, and cosine similarity suffer from their fair share of limitations. For instance, they ignore the ordering of topics which may carry information on relative emphases in a document. Future researches may consider applying a multitude of research approaches, including qualitative and quantitative techniques which may complement each other.

Manual analysis avoids measurement errors that computer programs may make when working with non-standard formats (Cazier & Pfeiffer, 2015). As the AICPA and the IFAC, publishers of the U.S. and the international ethics codes

respectively, are not obligated to standardize their formats in their respective ethics code of conduct, differences driven by non-standard formats are bound to exist. Manual content analysis can help moderate misinterpretations caused by these differences in the formats of ethics codes. Future research can therefore extend and improve this study by adopting both computerized and manual content analyses.

Many countries adopt the international ethics code with minor modifications. Comparing these modifications enables the reader to focus on the unique characteristics of each IFAC member country's ethics code. Future research may investigate how the nature and extent of these modifications relate to the stage of economic development, business practices, and national culture.

8. REFERENCES

- AICPA. (2016). AICPA Code of Professional Conduct. Durham, North Carolina.
- AICPA. (undated). Tax Ethics & Professional Standards Guidance. <https://www.aicpa.org/interestareas/tax/resources/standardsethics/taxethicsguidance.html> Assessed 11 August 2020.
- Allen, C. (2010). Comparing the Ethics Codes: AICPA and IFAC: Institute's Efforts Focus on Codifying and Aligning Rules with International Standards. *Journal of Accountancy*, 210(4), 24.
- Allen, C. R. (2018). A New Take on Ethics and Independence: A Refreshed International Code Emphasizes 3 Key Objectives for Professional Accountants. *Journal of Accountancy*, 226(6), 44.
- Anders, S. B. (2018). Ethics resources: IFAC and IESBA. *The CPA Journal*, 88(3), 72-73.
- Audi, R., Loughran, T., & McDonald, B. (2016). Trust, but verify: MD&A language and the role of trust in corporate culture. *Journal of Business Ethics*, 139(3), 551-561. <https://doi.org/10.1007%2Fs10551-015-2659-4>
- Ball, C., Hoberg, G. & Maksimovic, V. (2015). Disclosure, Business Change and Earnings Quality. Available at SSRN: <https://ssrn.com/abstract=2260371> Accessed 11 August 2020. <http://dx.doi.org/10.2139/ssrn.2260371>
- Bobek, D. D., Hageman, A. M., & Hatfield, R. C. (2010). The role of client advocacy in the development of tax professionals' advice. *Journal of the American Taxation Association*, 32(1), 25-51. <https://doi.org/10.2308%2Fjata.2010.32.1.25>

- Boskou, G., Kirkos, E., & Spathis, C. (2019). Classifying internal audit quality using textual analysis: the case of auditor selection. *Managerial Auditing Journal*, 34(8), 924-950. <https://doi.org/10.1108/maj-01-2018-1785>
- Bowne-Anderson, H. (2018). What data scientists really do, according to 35 data scientists? *Harvard Business Review*. <https://hbr.org/2018/08/what-data-scientists-really-do-according-to-35-data-scientists> Accessed 11 August 2020.
- Brown, S. V., & Tucker, J. W. (2011). Large-sample evidence on firms' year-over-year MD&A modifications. *Journal of Accounting Research*, 49(2), 309-346. <https://doi.org/10.1111/j.1475-679x.2010.00396.x>
- Calhoun, C. H., Oliverio, M. E., & Wolitzer, P. (1998). *Ethics and the CPA: building trust and value-added services*. NY: John Wiley & Sons.
- Cardona, R. J., Rezaee, Z., Rivera-Ortiz, W., & Vega-Vilca, J. C. (2019). Regulatory Enforcement of Accounting Ethics in Puerto Rico. *Journal of Business Ethics*, 155(3), 1-14. <https://doi.org/10.1007/s10551-019-04137-4>
- Cazier, R. A., & Pfeiffer, R. J. (2015). Why are 10-K filings so long? *Accounting Horizons*, 30(1), 1-21. <https://doi.org/10.2308/acch-51240>
- Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies*, 27(5), 1367-1403. <https://doi.org/10.1093/rfs/hhu001>
- Cheung, R. W. Y., & Agrawal, R. K. (2018). Ethics, accounting education and recruitment in Hong Kong. *International Journal of Accounting and Finance*, 8(3), 228-244. <https://doi.org/10.1504/ijaf.2018.097145>
- Christensen, A., Cote, J., & Latham, C.K. (2018). Developing ethical confidence: The impact of action-oriented ethics instruction in an accounting curriculum. *Journal of Business Ethics*, 153(4), 1157-1175. <https://doi.org/10.1007/s10551-016-3411-4>
- Clements, C., Neill, J. D., & Stovall, O. S. (2009a). An analysis of international accounting codes of conduct. *Journal of Business Ethics*, 87(1), 173-183. <https://doi.org/10.1007/s10551-008-9811-3>
- Clements, C. E., Neill, J. D., & Stovall, O.S. (2009b). The impact of cultural differences on the convergence of international accounting codes of ethics. *Journal of Business Ethics*, 90(3), 383-391. <https://doi.org/10.1007/s10551-010-0417-1>
- Cohen, J. R., & Pant, L.W. (1991). Beyond bean counting: Establishing high ethical standards in the public accounting profession. *Journal of Business Ethics*, 10(1), 45-56. <https://doi.org/10.1007/bf00383692>

Cohen, J. R., Pant, L. W., & Sharp, D.J. (1992). Cultural and socioeconomic constraints on international codes of ethics: Lessons from accounting. *Journal of Business Ethics*, 11(9), 687-700. <https://doi.org/10.1007/bf01686349>

Cohen, J. R., Pant, L. W., & Sharp, D. J. (1993). Culture-based ethical conflicts confronting multinational accounting firms. *Accounting Horizons*, 7(3), 1. <https://search.proquest.com/openview/433d97278a907bf1bc3b0967ff6f0265/1?pq-origsite=gscholar&cbl=3330>

Cohen, J. R., Pant, L. W., & Sharp, D. J. (1996). A methodological note on cross-cultural accounting ethics research. *The International Journal of Accounting*, 31(1), 55-66. [https://doi.org/10.1016/s0020-7063\(96\)90013-8](https://doi.org/10.1016/s0020-7063(96)90013-8)

Cooper, B. J., Leung, P., Dellaportas, S., Jackling, B., & Wong, G. (2008). Ethics education for accounting students—A toolkit approach. *Accounting Education: An International Journal*, 17(4), 405-430. <https://doi.org/10.1080/09639280802436681>

Curtis, G. (2008). The financial crisis and the collapse of ethical behavior. Greycourt White Paper, 44, 1-18.

Dolch, E. W. (1949). The use of vocabulary lists in predicting readability and in developing reading materials. *Elementary English*, 26(3), 142-177.

Erkan, Gunes. (2004). The University of Michigan in Novelty 2004. <https://trec.nist.gov/pubs/trec13/papers/umichigan.novelty.pdf> Accessed 11 August 2020.

Ertugrul, M., Lei, J., Qiu, J., & Wan, C. (2017). Annual report readability, tone ambiguity, and the cost of borrowing. *Journal of Financial and Quantitative Analysis*, 52(2), 811-836. <https://doi.org/10.1017/s0022109017000187>

Fatemi, D., Hasseldine, J., & Hite, P. (2020). The Influence of ethical codes of conduct on professionalism in tax practice. *Journal of Business Ethics*, 164(1) 133-149. <https://doi.org/10.1007/s10551-018-4081-1>

Fischer, M., & Rosenzweig, K. (1995). Attitudes of students and accounting practitioners concerning ethical acceptability of earnings management. *Journal of business ethics*, 14(6), 433-444. <https://doi.org/10.1007/bf00872085>

Flory, S. M., Phillips Jr, T. J., Reidenbach, R. E., & Robin, D. P. (1992). A multidimensional analysis of selected ethical issues in accounting. *The Accounting Review*, 67(2), 284-302. https://www.jstor.org/stable/247725?seq=1#metadata_info_tab_contents

Gandy, L. M., Gumm, J., Fertig, B., Thessen, A., Kennish, M. J., Chavan, S., Marchionni, L., Xia, X., Shankrit, S., & Fertig, E. J. (2017). Synthesizer: Expediting synthesis studies from context-free data with information retrieval

techniques. *PLoS one*, 12(4), e0175860. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0175860> Accessed 11 August 2020. <https://doi.org/10.1371/journal.pone.0175860>

Goel, S. (2019). Ethical Accounting: The Driver in Recovering Markets. In Bilgin, M.H. (eds.), *Eurasian Business Perspectives* (pp. 95-106). Springer, Cham. https://doi.org/10.1007/978-3-030-11872-3_7

Hanley, K. W., & G. Hoberg. (2012). Litigation risk, strategic disclosure and the underpricing of initial public offerings. *Journal of Financial Economics*, 103 (2): 235-254. <https://doi.org/10.1016/j.jfineco.2011.09.006>

Heston, S. L., & Sinha, N. R. (2015). News versus sentiment: Predicting stock returns from news stories. FEDS Working Paper No. 2016-048. <https://www.federalreserve.gov/econresdata/feds/2016/files/2016048pap.pdf> Accessed 11 August 2020. <http://dx.doi.org/10.17016/FEDS.2016.048>

Hofstede, G. (1980). *Culture's Consequences*. Beverly Hills, CA: Sage Publications.

Hofstede, G. (1991). *Cultures and Organizations: Software of the Mind*. New York: McGraw Hill.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage Publications.

Hofstede, G., Hofstede, G.J., & Minkov, M. (2010). *Cultures and Organizations: Software of the Mind. Revised and expanded (3rd ed.)*. New York: McGraw-Hill.

Hu, Y., Chen, S., & Wang, J. (2018). Managerial Humanistic Attention and CSR: Do Firm Characteristics Matter? *Sustainability*, 10(11), 4029. <https://doi.org/10.3390/su10114029>

International Federation of Accountants (IFAC) (2018). *Handbook of the International Code of Ethics for Professional Accountants*. New York, New York.

Jung, Y. (2018). Corporate Strategies, Human Capital, and the Management of High-Tech Firms. <https://www.semanticscholar.org/paper/Corporate-Strategies-%2C-Human-Capital-%2C-and-the-of-Jung/78e3ed6a13cec28da03e13c1806533015c715ce8> Accessed 11 August 2020.

Kelly, K., & Murphy, P. R. (2019). Reducing Accounting Aggressiveness with General Ethical Norms and Decision Structure. *Journal of Business Ethics*, 1-17. <https://doi.org/10.1007/s10551-019-04290-w>.

Kothari, S., Li, X., & Short, J. (2009). The effect of disclosure by management, analysts, and financial press on cost of capital, return volatility, and analyst

- forecasts: A study using content analysis. *The Accounting Review*, 84 (5), 1639-1670. <https://doi.org/10.2308/accr.2009.84.5.1639>
- Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2-3), 110-135. <https://doi.org/10.1016/j.jacceco.2015.09.002>
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3), 221-247. <https://doi.org/10.1016/j.jacceco.2008.02.003>
- Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting and Economics*, 63(1), 1-25. <https://doi.org/10.1016/j.jacceco.2016.09.002>
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4), 1643-1671. <https://doi.org/10.1111/jofi.12162>
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187-1230. <https://doi.org/10.1111/1475-679x.12123>
- Lundholm, R. J., Rogo, R., & Zhang, J. L. (2014). Restoring the tower of Babel: How foreign firms communicate with US investors. *The Accounting Review*, 89(4), 1453-1485. <https://doi.org/10.2308/accr-50725>
- Manning, C. D., & Schütze, H. (1999). Foundations of statistical natural language processing. Cambridge, Mass: MIT press.
- Martinov-Bennie, N., & Pflugrath, G. (2009). The strength of an accounting firm's ethical environment and the quality of auditors' judgments. *Journal of Business Ethics*, 87(2), 237-253. <https://doi.org/10.1007/s10551-008-9882-1>
- Matsumoto, D., Pronk, M., & Rolelofsen, E. (2011). What makes conference calls useful? The information content of managers' presentations and analysts' discussion sessions. *The Accounting Review*, 86(4), 1383-1414. <https://doi.org/10.2308/accr-10034>
- Mayew, W. J., & Venkatachalam, M. (2012). The power of voice: Managerial affective states and future firm performance. *The Journal of Finance*, 67(1), 1-43. <https://doi.org/10.1111/j.1540-6261.2011.01705.x>

- Melé, D. (2005). Ethical education in accounting: Integrating rules, values and virtues. *Journal of Business Ethics*, 57(1), 97-109. <https://doi.org/10.1007/s10551-004-3829-y>
- Miller, W. F., & Shawver, T.J. (2018). An exploration of the state of ethics in UK accounting education. *Journal of Business Ethics*, 153(4), 1109-1120. <https://doi.org/10.1007/s10551-016-3396-z>
- Pergola, T. M., & Walters, L. M. (2017). Ethics in the accounting curriculum: A model for an enhanced accounting ethics course. *Journal of Business Ethics Education*, 14, 199-228. <https://doi.org/10.5840/jbee20171410>
- Pieters, M. (2019). International code of ethics for professional accountants: A behind the scenes look at the eCode. *Australian Accounting Review*, 29(3), 595-598. <https://doi.org/10.1111/auar.12295>
- Ramage, D., Rafferty, A. N., & Manning, C. D. (2009, August). Random walks for text semantic similarity. In Proceedings of the 2009 workshop on graph-based methods for natural language processing (pp. 23-31). Association for Computational Linguistics. <https://nlp.stanford.edu/pubs/wordwalk-textgraphs09.pdf> Accessed 11 August 2020. <https://doi.org/10.3115/1708124.1708131>
- Reinstein, A., Churyk, N. T., Taylor, E. Z., & Williams, P. F. (2019). Using Accounting Department Advisory Councils and Higher Quality Continuing Education Requirements to Improve the Accounting Profession's Ethical Reasoning Skills. *Advances in Accounting Education: Teaching and Curriculum Innovations* (Advances in Accounting Education, Vol. 23), Emerald Publishing Limited, 177-193. <https://doi.org/10.1108/S1085-462220190000023012>
- RxNLP. (2019). What is Text Similarity API? <https://rxnlp.com/api-reference/text-similarity-api-reference/#.XMD7juhKi71> Accessed 11 August 2020.
- SEC. (2007). Speech by SEC chairman: Closing remarks to the second annual corporate governance summit. <https://www.sec.gov/news/speech/2007/spch032307cc.htm> Accessed 11 August 2020.
- Shawver, T. J., & Miller, W. F. (2017). Moral intensity revisited: Measuring the benefit of accounting ethics interventions. *Journal of Business Ethics*, 141(3), 587-603. <https://doi.org/10.1007/s10551-015-2711-4>
- Sonpar, K., & Golden-Biddle, K. (2008). Using content analysis to elaborate adolescent theories of organization. *Organizational Research Methods*, 11(4), 795-814. <https://doi.org/10.1177/1094428106297804>

Spalding, A. D., & Oddo, A. (2011). It's time for principles-based accounting ethics. *Journal of Business Ethics*, 99(1), 49-59. <https://doi.org/10.1007/s10551-011-1166-5>

Stahl, S. A. (2003). Vocabulary and readability: How knowing word meanings affects comprehension. *Topics in Language Disorders*, 23(3), 241-247. <https://doi.org/10.1097/00011363-200307000-00009>

Stanga, K. G., & Turpen, R. A. (1991). Ethical judgments on selected accounting issues: An empirical study. *Journal of Business Ethics*, 10(10), 739-747. <https://doi.org/10.1007/bf00705708>

Tetlock, C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63, 1437-1467. <https://doi.org/10.1111/j.1540-6261.2008.01362.x>

Tweedie, D., Dyball, M. C., Hazelton, J., & Wright, S. (2013). Teaching global ethical standards: A case and strategy for broadening the accounting ethics curriculum. *Journal of Business Ethics*, 115(1), 1-15. <https://doi.org/10.1007/s10551-012-1364-9>

Xu, Q., Fernando, G. D., & Tam, K. (2018). Executive age and the readability of financial reports. *Advances in Accounting*, 43, 70-81. <https://doi.org/10.1016/j.adiac.2018.09.004>

Yang, J., & Park, K.M. (2017). The Effects of Performance Feedback, Past Temporal Orientation, and Interaction on R&D Intensity. In *Academy of Management Proceedings 2017(1)*, 17526. Briarcliff Manor, NY 10510: Academy of Management. <https://doi.org/10.5465/ambpp.2017.17526abstract>