

# **FARS Database Searching: Providing Potential Search Terms to Students**

**Margaret R. Garnsey\***  
Siena College

**John O'Neill**  
Siena College

**Leonard Stokes**  
Siena College

## **Abstract**

This research examines the effect of providing students with a list of potentially relevant terms taken directly from the FARS (Financial Accounting Research System) database on their ability to locate the answer to a research question. Over a period of two years, 127 students were tested, in class, on the use of FARS as a research tool. Students were given two research questions: one included only an accounting dilemma and the other included a different accounting dilemma and a list of terms generated by the use of agglomerative clustering. The study found that when student were given a list of potential search terms the proportion of correct answers increased for one of the accounting dilemmas. It was also found that the proportion of "Off the Mark" answers decreased for both dilemmas when a list was available. These significant findings have an impact on educators and developers of the CPA professional exam. Providing a list of available terms, more closely resembles what is available in practice where team work is important.

## **Introduction and Background**

The objective of this study was to determine if suggesting possible query terms, which were taken directly from the FARS (Financial Accounting Research System) database, improved student search results. This is an important endeavor because if search accuracy is improved, it will be easier for stakeholders to find necessary information more quickly and efficiently. It further develops research by Garnsey (2006), who showed that statistical methods can be used to form groups of related terms using the vocabulary taken directly from the FARS database and by Garnsey & Hotaling (2007), who established that accounting educators had greater concept recall when presented with a group of terms created using those statistical methods. Allowing students to select which terms they think are relevant to a particular information need should avoid the problem of returning large result sets that may be the consequence of using automatic query expansion via the Folio Views synonyms.

---

\* Corresponding author, contact to obtain data used in the study

FARS is an important research tool for the professional and the student. The AICPA Uniform CPA Examination, Examination Content Specifications identifies one of the skills required of entry level accountants as the ability to "perform research and analyze information" (AICPA, 2002). FARS research requirements using simulations have been included in the computerized CPA exam since April, 2004. The definition of simulations includes "the ability to use electronic tools such as databases."

Perceived shortcomings in the authoritative accounting literature are well known. Gangolly, Hedley, & Wong (1991), stated "The accounting standards ...exhibit ambiguity, semantic richness, and extremely complicated rules for inference..." They proposed using an object-oriented language to represent the rules in accounting in order to develop a knowledge base. This would require an in-depth semantic analysis of the accounting standards which has not been undertaken at the present time. More recently Fisher (2003), states: " Financial accounting standards, as currently drafted, lack a consistent and sound structural arrangement and are incompatible with digital storage and retrieval methods." She makes suggestions for restructuring Financial Accounting Standards to facilitate the application of digital techniques as well as to enhance the understandability of standards. Her proposal includes the use of mark-up languages to formally describe and standardize the structure of accounting standards. In September 2004, the Financial Accounting Foundation (FAF) trustees approved the FASB's Codification and Retrieval project, which is designed to develop an authoritative, searchable retrieval system for a single, topically organized database incorporating all United States Generally Accepted Accounting Practices (GAAP) (FASB, 2004). In 2008, the codification project had just entered its verification stage.

This paper is organized as follows. The next section reviews prior research on the cognitive characteristics of accountants and vocabulary problems related to information retrieval. Understanding these characteristics allows the reader to understand potential difficulties inexperienced accountants may have in retrieving relevant information from GAAP. This section also gives an overview of *Folio Views*, the software interface of the FARS database, and explains how its query expansion features may actually degrade query results. The following section provides the development of the hypotheses that were tested and describes the methodology used to conduct the experiment. It reviews prior research related to the development of the hypotheses. It then describes the experiment which was conducted over a two year period. The last sections present the results and discussion, including suggestions for future research.

## Theory

Recalling that the objective of the study is to determine whether or not a list of semantic terms will improve accuracy in search results, it is important to understand the cognitive processes that students undergo when faced with problems contained in a simulation. Secondly, it is important to recognize the vocabulary problems that can occur as impediments to proper query development.

Fortunately, many accountants have been involved in various studies over the years to determine their cognitive capabilities and intellectual characteristics. In addition, accounting firms have been concerned about these characteristics for individuals entering the profession. As such, they have used the CPA exam as an agent to test for these cognitive characteristics.

In reviewing literature on the cognitive characteristics of accountants Ho & Rodgers (1993) used Kogan's Taxonomy. This taxonomy breaks cognitive issues into styles, abilities, and strategies. *Cognitive styles* remain consistent over time and are distinctive ways of acquiring, storing, retrieving, and transforming information. *Cognitive abilities* relate to knowledge encoding and retrieval and skill level. These abilities certainly change throughout an individual's lifetime. *Cognitive strategies* focus on interactions between cognitive style, cognitive abilities and the individual's environment. They are affected by task requirements, situational constraints, and problem contents.

Ho & Rodgers reviewed a variety of studies that they classified as evaluating cognitive style. Many of these studies used, for example the Myers-Briggs Type Indicator (MBTI), to establish an individual's psychological preferences

in perceiving the world and making decisions. Studies that Ho & Rodgers classified as examining cognitive strategies used for example, the intolerance of ambiguity. The level of intolerance of ambiguity can affect the amount and nature of the information received.

Ho and Rodgers (1993) state that “cognitive abilities cannot be well understood without reference to knowledge.” They point out that studies aimed at problem solving demonstrate that “... experience enhances the representation of the problem and, through feedback and learning, turns novices into relatively more expert decision makers.”

They mention previous research on problem solving that suggests decision makers in complex situations deliberately simplify the choice mechanism. For example, “... people faced with a complex decision task may employ strategies to eliminate some of the alternatives” (Ho & Rodgers, 1993). They maintain “cognitive strategies may be refined with practice... and thus are more amenable to change through training.” This suggests that persons with limited experience in problem solving would not have well-developed cognitive strategies.

Therefore, persons taking the CPA exam may have trouble performing on simulation questions since they are typically inexperienced professionals. For inexperienced users, one of the most frequent difficulties is determining relevant query terms. In practice these persons are typically on a team of professional auditors who are more experienced in developing research strategy. On the CPA exam, they are confronted (via simulation) with the task of using technical research to reach professional judgments in an *individual* setting. This then motivates the question of whether poor performance on a simulation is due to poor query development leading to inaccurate retrieval information or the inability to recognize retrieved information as relevant to the research simulation.

Related to the questions surrounding poor performance on simulations, prior research has shown the retrieval of information involves three stages (Foltz, 1991): generating retrieval cues, using them to retrieve information, and evaluating whether the retrieved information is what is desired. All information retrieval systems (including the FARS database) depend on the user entering some form of query into the system. The query terms are then used to determine which items in the system are relevant to the query. The system determines relevance by some method for matching query terms to document terms. Frequently, all relevant documents are not retrieved by these methods. Based on their research with full-text document-retrieval systems, (Blair & Maron, 1985) state, “A full-text retrieval system...places the user in the position of having to find an impossibly difficult combination of search terms.”

The first step of the retrieval process, generating retrieval cues, is crucial in obtaining relevant results. This may be difficult if the problem that gave rise to the information need is not well defined. It is even more difficult for users who are unfamiliar with the topic and have a limited vocabulary to describe their need. Finally, users may not have knowledge of the database vocabulary. Inexperienced users of financial information are frequently unfamiliar with the variety of terms which can represent a single concept in accounting. In a term association test given to seven accounting teachers, Garnsey & Hotaling (2007) found that for 13 of 30 randomly selected accounting terms, at least one of the seven teachers could not generate any related terms. They further found that the number of terms with no response declined significantly when the subjects were provided with a list of potentially related terms.

Furnas, Landauer, Gomez, & Dumais (1987) used five different application domains and required people from different groups to name selected objects from them. They found that when access to an item was provided by any single one of the names chosen, untutored people failed to access the item 80-90% of the time on their first attempt. Their research found that the probability of two people using the same word to describe a given object averaged less than 20%. Chen & Lynch (1992) found that “the consistency levels of terms selected by different subjects were between 8% and 10%.” This research suggests a need to provide inexperienced users with clues that will increase their chances of locating relevant information in electronic databases.

FARS is an example of just such an electronic database. It uses the generic product Folio Views which combines a search engine supporting Boolean, wildcard, structured, and natural language searching with a user interface. Because Folio Views is a generic interface some of its features may degrade search performance rather than improve

it; for example, the Query dialog "expands searches automatically by applying the word stem operator to the terms" (Folio Views, 2003). In accounting, terms with the same stem can have very different meanings (e.g. warrant, warranty), therefore, the automatic expansion can decrease search precision. Folio Views also allows the user to include synonyms in a query term by adding a "\$" after the term. For example, if bond\$ is used in a query the hit list contains matches with the terms given in table 1. While some of the terms may be appropriate, the majority would be considered irrelevant for a search on the concept that accountants mean when the term "bond" is used. This means the search results will be less precise because the Folio Views synonyms are generic rather than accounting specific. Our research tests whether providing users with terms taken directly from the FARS database and allowing them to choose which terms to use in augmenting queries results in more successful searches. The following section briefly discusses issues in information retrieval and then presents the hypothesis used in evaluating our research results.

### Hypothesis Development

Anecdotal evidence and research on searching the Internet indicates "most users will only have the time or patience to evaluate the top ten or twenty Web pages that are returned..." (Clarke, 2000). Using the above example, the number of hits expands from 125 to 3,455 when including the synonyms of bond. However, most of the additional hits pertain to the wrong concepts. If only a limited number of those hits are reviewed, query expansion could *decrease* the chance of finding the relevant information rather than increase it.

Retrieval thesauri are a common means of improving retrieval. They have traditionally been prepared manually in a time consuming and expensive process. Price Waterhouse & Co. (1974), one of the then "big six" accounting firms, published an accounting thesaurus for internal use. However, due to the evolution of accounting language, many terms that are common today (e.g. functional currency, derivatives) are not included in it. Muddamalle (1998) found that only 45% of request terms matched terms available in the *Microthesaurus of Soil Mechanics Terms* which was published in 1974. That research implies that even if the Price Waterhouse thesaurus was widely available, because of its age, it would have limited usefulness.

Individuals requesting information typically use a limited vocabulary to express their information need. Because of the variety of terms that can convey the same concept searchers may not express their request in vocabulary that matches the relevant information in the database. Such a query would have a relatively large semantic distance. For example, an individual interested in method of estimating accounts that people will not pay, could use terms such as "bad debts," uncollectible accounts," or "doubtful accounts." Their search results depend on the actual terms used in FARS. If the users depend on the automatic query expansion of the FARS database using synonyms, then they will obtain more hits. However, those hits will contain a higher proportion of irrelevant results because the synonym list is not accounting specific and all terms on the list are automatically added to the query. This makes it less likely they will be able to identify the correct information to answer their information need.

Borthick, Bowen, Jones, & Tse (2001) investigated the effect of ambiguity on query development in relational databases. They took information requests and varied the semantic distance between the information request and the expression in the query language that would produce the desired information. In the context of their research, *semantic distance* is defined as "the distance between the information users want and the expression in the query interface language that will produce that information." They found that accuracy and efficiency were inversely related to the ambiguity of information requests.

We present users with an accounting specific list of terms obtained directly from FARS decreasing the semantic distance. The lists were taken from clusters of related terms derived by Garnsey (2001). That research combined Latent Semantic indexing (LSI) and agglomerative clustering to group together related accounting concepts using the vocabulary taken directly from FARS. LSI is a method of two mode factor analysis that infers relationships between terms that are used in the same context, even if the terms have never been used together in the same document. Agglomerative clustering begins with each term as a separate cluster and successively merges the two

most similar clusters until all terms have been merged. Garnsey & Hotaling (2007) found that clusters generated in this manner enabled subjects to identify significantly more relevant terms than they were able to from memory alone.

For the first question clusters containing the term "contingencies" were used. Fourteen students (who were not given a list) used either the singular or plural contingencies, verifying that this was a term students would know and use. Table 2 shows the cluster terms and their potential for leading students to the correct paragraph(s).

For the second question, clusters containing the term "interest rates" were used. The term "interest rate" was used in the fact pattern given to the students. Table 3 shows the cluster terms and their potential for leading students to the correct paragraph(s).

By allowing users to select which terms to use they are able to select items that they feel are relevant to their particular information need. This makes it more likely that the results of the query will contain fewer irrelevant results and enable them to identify the information applicable to their request leading to the following hypothesis:

- 1) When users are given a list of terms taken directly from FARS the number of correct responses will increase ( $H_1$ )
- 2) When users are given a list of terms taken directly from FARS the number of responses Off the Mark will decrease ( $H_2$ )

## Methodology

To test the hypotheses, the following two-stage study was designed for students in the second semester of an intermediate accounting sequence, prior to any courses including business combinations or international accounting. Its participants had previously been exposed to many GAAP concepts including topics relevant to the research questions used in this study. They were also introduced to the use of FARS in the classroom and had been informed that there would be a FARS research question on their final exam. Furthermore, they were also given a homework assignment which required its usage, although no formal advanced training in the use of search techniques had been provided during their classroom experience.

These students were asked to identify the correct paragraph in FARS for two research questions (Appendix A) and complete a short survey (Appendix B) at the end of one class period. Students were given a list of possible query terms, using the clusters developed by Garnsey (2006) and discussed in the previous section, for one research question and were asked to complete the other research inquiry without a list of terms. For this study, clusters of terms less than ten words in size were used due to the time restrictions that students had to complete the assignment. This 25 minute time limit to obtain the results was decided upon for a variety of reasons including space limitations, licensing limitations and a time period similar to CPA exam constraints. Question 1h designates that students received the aforementioned list of terms in searching for their answer to question 1. Question 2h designates that students received their list of terms to assist them in searching for their answer to question 2. Therefore an individual student would have (in order) received questions 1 and 2h or Questions 1h and 2 to perform their searches.

Each of the questions had a predetermined correct response. The professor assessing the responses identified one or more paragraphs that would qualify as the correct (C) research result (see Teaching Note Appendix C), if the student could locate it. The professor did not make any judgment about whether or not the student would have interpreted the proper paragraph(s) correctly or incorrectly. The student designated search results were placed in one of three categories: "Correct" (C), "Needs More Work" (NMW), or "Off the Mark" (OTM). In the category NMW, the student had located the correct SFAS or APBO but was not within what the professor had determined to be the proper paragraphs. The professor assessing the responses determined that a student in the NMW category would have been able to locate the correct paragraphs if given more time. The category OTM indicated that the student had not located the correct SFAS or APBO.

## Results

The results of this study have been separated into three parts. The first subsection explains the outcome of the experiment in descriptive terms, while the second presents the results of the tests of hypotheses and what inferences we may draw from them. The last subsection presents a summary of student comments to a short survey. It supports the results in the previous two sections.

### *Overview of Search Results*

For the first question, all of the students without a list who used the term contingencies were able to identify the correct SFAS and some identified the correct paragraph. Other students (without the list) who used contingencies in conjunction with other terms were also able to identify the correct SFAS. Of the students who were not given a list and who were not able to identify the correct SFAS, very few used any list terms and, for the most part, when list terms were used it was in conjunction with other terms. FolioViews “ands” terms together, meaning all terms used must appear in the result. Therefore, the use of additional terms with contingencies could result in relevant results being eliminated. It is interesting to note that almost 30% of students misinterpreted the question as an environmental problem rather than a contingent liability when they were not given a list, as opposed to approximately 11% when they were given a list.

For the second question, the majority of students who were not given a list and identified the correct APBO(s) used at least one term from the clusters. Of the students who were not given a list and who were not able to identify the correct APBO(s), about half used the list term “bonds” in conjunction with other terms and several used the singular of the term “bond” in conjunction with other terms. Several of these students also used the term “interest rate.” It is not clear why these students were unable to identify a correct APBO. One theory is that since the terms were used in conjunction with other terms, the correct results may have been screened out when the search became too narrow.

As mentioned above, research on search behavior and the Internet has indicated that most users will only evaluate the top ten to twenty pages in a search (Clarke, 2000). The same could be inferred for other types of key word searches. This leads to a competing theory that because the search terms used for the second question were more general than the terms for the first question (see Table 4 for cluster terms which gave correct APBOs or SFASs) students did not review all of the search results. This is especially true because the advanced search results are presented chronologically in FARS.

### *Hypothesis Testing*

When separated by question, it should be noted that each sample is independent since the two sets of responders are disjoint. Additionally, each sample is sufficiently large that the sampling distribution of the mean is approximately normally distributed ( $n \geq 63$ ) except when the categorical proportion is small (e.g.  $p < 7\%$  or  $p > 93\%$ ). Therefore, we used the one-tailed Z-test for the difference between two population proportions and can formally restate the hypotheses as follows.

$H_1$ : When users are given a list of terms taken directly from FARS the number of correct responses will increase.

$$H_o : \pi_{CH} \leq \pi_C$$

$$H_a : \pi_{CH} > \pi_C$$

$H_2$ : When users are given a list of terms taken directly from FARS the number of responses deemed “off the mark” (OTM) will decrease.

$$H_o : \pi_{OH} > \pi_O$$

$$H_a : \pi_{OH} \leq \pi_O$$

For the 127 students participating in the study, the results are summarized in Table 5.

When students were given a list of terms, the initial observation is that, for both questions, the number of Correct responses increased and the number of Off the Mark (OTM) responses decreased. Secondly, the number of Needs More Work (NMW) responses increase for each question, although the nature of this category and the presentation of aggregate data make it difficult to capture any changes in the category. The results of the hypothesis tests as performed on each question are stated in Table 6.

For question 1, we see a significant increase in Correct Responses. For question 2, the increase of correct responses is not statistically significant. For **both** questions, we see a significant decrease in the number of OTM responses. These results indicate that providing a list of potentially relevant terms increases the accuracy of student searching.

### *Survey Results*

Results from the short survey indicated a student attitude that agreed with the test results. About seventy-five percent of the participants reported that the list of potential terms was helpful in the search and that having the potential terms allowed them to obtain what they felt was a better or more complete answer. This proportion increases to eighty percent when one examines students who answered at least one question correctly. Slightly less than half of the participants thought having the potential terms made the search shorter.

The most frequently occurring comments were about the time constraints. A typical general comment was “It’s not enough time, to do research [we] need more time.” The following comment was made in response to the question of whether the list of potential terms allowed the student to obtain what they felt was a better or more complete response: “Possibly, but didn’t have enough time to thoroughly read all of the results.” Other comments included:

- ❖ “Typing in potential terms helped a lot but a lot of the time they led you to the FARS paragraphs that were not related.”
- ❖ “It confirmed the answers that I worked on. This exercise gives a person a way to search for information on topic that is unclear.”
- ❖ “FARS is way too in depth to find such particular problems with a quick amount of time.”
- ❖ “...However, FARS, doesn’t give the search results you want with one term. The best results I found were with multiple terms, adding terms that were not given.”
- ❖ “Made the selection process faster. Process as a whole still very long and time consuming.”
- ❖ “I found information that was relevant to the search terms but not the scenario situation.”
- ❖ “Good exercise but the computer program is too slow to complete the exercise in timely manner.”

## Discussion and Future Research

Terms suggested by the clusters clearly reduced the number of students who were unable to locate the pronouncement containing the answer to a research question. The tests of hypotheses produced statistically significant results. The number of students who were incorrect decreased significantly. The use of accounting specific terms taken directly from the FARS database provides students with a means of reducing the semantic distance between their information need and their query to the database. It may be possible to use statistically related terms as a first step in addressing the AICPA Special Task Force's recommendation for implementation of a system to facilitate access, and to improve the understanding of the professional literature (O'Dell, 1998).

The percent of students, without a list, whose responses were "Off the Mark" (Q1, 59% and Q2, 67%) is comparable to the percentage of students failing (63%) the FAR section of the Uniform CPA exam in 2004 (NASBA, 2005). The fact that many student responses were "Off the Mark" may have been the result of FARS coming up with a response that was not appropriate and the student either not knowing this was an inappropriate response, or not having the time or research ability to get closer. The research techniques of the students were shown to be significantly equivalent, yet the research results when using FARS can be difficult to determine. Several students commented about the difficulty in using FARS, and compared it unfavorably with RIA, the package they use for tax research. Within the experimental design, it was difficult to separate the behavioral issues from the quantitative results. More research on how individuals react to the myriad of search results produced by querying an electronic database is important. Additional research is needed to see if clusters of related terms can be generated more precisely. Clusters were generated using a term-document matrix. Many documents in the FARS database cover several subjects. It may be appropriate to subdivide these documents into sections or paragraphs. In addition, the number of dimensions used for the latent semantic indexing was chosen empirically and may not have been the optimum. Preliminary research in determining the optimal number of factors has been initiated (Ding, 1999). This research should be confirmed and extended. Finally, by including appropriate collocations (terms appearing next to each other) determined directly from the corpus, it may be possible to significantly expand the number of meaningful phrases included in the vocabulary.

One area of future research is to try and evaluate the ramifications of time restrictions. Many students commented about having insufficient time to complete the assignment. In several cases, they felt their answer would have been more complete if given more time. Several students commented that their searches went in circles. Although it is likely that the majority of time was used to evaluate the returned items, it is also possible that more time is needed to generate appropriate terms for queries. It may be possible to break down the search process to evaluate where students are spending their time.

Another area of future study is to determine whether or not receiving a list first improves search results in subsequent tasks. Because of the setup of this study, it is unknown whether receiving a list of related terms influences a subject's subsequent cognitive strategy.

Student search strategies also need to be researched. Very few students indicated that they used a phrase (by placing the terms in quotes). Many students used terms which would not differentiate content. For example, the term "accounting" occurs over 7,000 times in the FARS database; similarly, "disclosure" appears over 1,500 times. However, a fair number of students used these terms in their searches. This implies that instructing students in efficient search techniques would improve their performance. Students could also be instructed on how FARS displays search results; chronologically by date of issuance and out of context with the statement in which the paragraph containing the search term(s) appears. This facet of research may also provide a partial answer to the question of why students were off the mark so frequently. For instance, if a student knows how to use proper search techniques then if that student knew the answer to the problem he/she would be more likely to locate the correct SFAS and improve their off the mark result to a correct one.

## References

- AICPA. (2002, 10/19/05). *Uniform CPA Examination: Examination Content Specifications*, [http://www.cpa-exam.org/download/CPA\\_Exam\\_CSOs\\_revised\\_10\\_05.pdf](http://www.cpa-exam.org/download/CPA_Exam_CSOs_revised_10_05.pdf) (Accessed: 1/24/2009).
- Arthur Anderson & Co., e. a. (1989). *Perspectives on Education Capabilities for Success in the Accounting Profession*.
- Blair, D. C., & Maron, M. E. (1985). An Evaluation of Retrieval Effectiveness for a Full-Test Document-Retrieval System. *Communications of the ACM*, 28(3), 289-299.
- Borthick, A. F., Bowen, P. L., Jones, D. R., & Tse, M. H. K. (2001). The effects of information request ambiguity and construct incongruence on query development. *Decision Support Systems*, 32, 3-25.
- Chen, H., & Lynch, K. J. (1992). Automatic Construction of Networks of Concepts Characterizing Document Databases. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(5), 885-902.
- Clarke, S. J. (2000). Search Engines for the World Wide Web: An Evaluation of Recent Developments. *Journal of Internet Cataloging*, 2(3/4), 88-91.
- Ding, C. H. Q. (1999). *A Similarity-Based Probability Model for Latent Semantic Indexing*. Paper presented at the 22nd annual international ACM SIGIR conference on Research and Development in Information Retrieval, Berkeley, Ca.
- FASB. (2004). Codification and Retrieval Project. *The FASB Report, September 30, 2004*, [http://www.fasb.org/articles&reports/fasb\\_embarks\\_tfr\\_sept\\_2004.pdf](http://www.fasb.org/articles&reports/fasb_embarks_tfr_sept_2004.pdf) (Accessed: 1/24/2009).
- Fisher, I. E. (2003). On the Structure of Financial Accounting Standards to Support Digital Representation, Storage and Retrieval. *Journal of Emerging Technologies in Accounting*, 1(1).
- Folio Views. (2003). *Folio 4 Help Infobase*.
- Foltz, P. W. (1991, August 7-10, 1991). *Modeling Human Memory Retrieval and Computer Information Retrieval: What can the two fields learn from each other?* Paper presented at the Thirteenth Annual Conference of the Cognitive Science Society, Chicago, Illinois.
- Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1987). The Vocabulary Problem in Human-System Communication. *Communications of the ACM*, 30(11), 964-971.
- Gangolly, J. S., Hedley, T. P., & Wong, C.-T. (1991). Semantic Knowledge Bases for Financial Accounting Standards. *Expert Systems with Applications*, 3, 117-128.
- Garnsey, M. R. (2006). Automatic Classification of Financial Accounting Concepts. *Journal of Emerging Technologies in Accounting*, 3, 21-39.
- Garnsey, M. R., & Hotaling, A. W. (2007). Improving Identification of Search Terms in the FARS Database. *Review of Business Information Systems*, 11(1), 45-55.

- Ho, J. L., & Rodgers, W. (1993). A Review of Accounting Research on Cognitive Characteristics. *Journal of Accounting Literature*, 12(1993), 101-130.
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley & Sons, Inc.
- Muddamalle, M.R. (1998). Natural Language versus Controlled Vocabulary in Information Retrieval: A Case Study in Soil Mechanics. *Journal of the American Society for Information Science*, 49(10), 881-887.
- NASBA (2005) *Candidate Performance on the Uniform CPA Examination: Reports and commentary on the 2004 CPA Examination*. Nashville National Association of State Boards of Accountancy
- O'Dell, J. H. (1998). Accounting & Auditing. *Pennsylvania CPA Journal*, 68(4), 6-7.
- Price Waterhouse & Co. (1974). *Thesaurus of Accounting and Auditing Terminology*. New York: Price Waterhouse & Co.
- Venables, W. N., & Ripley, B. D. (1997). *Modern Applied Statistics with S-Plus*. New York: Springer.
- Willett, P. (1988). Recent Trends in Hierarchic Document Clustering: A Critical Review. *Information Processing & Management*, 24(5), 577-597.

## Appendix A

### Question 1

Smithtown Manufacturers are your client. A toxic dumpsite has been found on their property during the year. Their management thinks they will have to clean it up. Their attorneys have issued an opinion that they are probably liable. The cost estimates for the clean-up range from \$250,000-\$500,000. Management want you to provide an opinion based upon GAAP on the proper accounting this year as well as any required disclosures to be made in this year's financial statements. Using FARS locate the paragraph(s) that need to be quoted. Provide the words you used to search and the result of the search. Please indicate which paragraphs you will quote.

### Question 2

Clifton Industries issued \$10,000,000 bonds for \$9,100,000. The Bonds pay 10% interest and will mature in 10 years. The controller wants to use the effective interest rate method and the CFO wants to use the straight line method to amortize the discount. You have been asked to explain what is the appropriate procedure to use to record the interest payments.

#### Terms used for Question 1h

The following terms may be, but are not required to be considered:

change in accounting estimate
contingencies
errors
litigation
loss contingencies
prior period adjustments
renegotiation
adjustments related to prior interim periods of the current fiscal year

#### Terms used for Question 2h

The following terms may be, but are not required to be considered:

banks
bonds
cash flow
interest rates
markets
maturity
notes
principal

## Appendix B

**STUDENT COMMENT FORM AFTER COMPLETING FARS RESEARCH**

What are your Scenario #'s \_\_\_\_\_

Please answer the following questions:

Was using a list of potential terms helpful in the search?                      Yes                      No

If you answered the above question **Yes**, please answer the following:

Did having potential terms make the search more complete?                      Yes                      No

(i.e.: you found information you may not have otherwise)

Did having potential terms make the search shorter?                      Yes                      No

Did having the potential terms allow you to obtain what you feel

is a better or more complete answer?

Please share any comments you have about this exercise:

If you answered **NO** please explain why you feel this way?

Please share any comments you have about this exercise:

Overall how would rate the use of having access to a list of potential terms?

Minimal value

Great Value

1

2

3

4

5

6

7

## Appendix C

### Teaching Note

This exercise was designed to allow students to have the opportunity to research using the FARS database with some level of comfort related to the topic area. Both “Contingencies” and Effective Interest Rates” are discussed during the semester. While individual pronouncements are mentioned during the semester there is no emphasis on the students memorizing technical pronouncement names, numbers or paragraphs.

During the exercise the students are expected to list the Pronouncement number and paragraph they are not expected to write a full analysis. As such I expect as correct a list of pronouncements and if the appropriate paragraphs are listed the response was considered to be “Correct”. If the student located the correct pronouncement but did not list the proper paragraph or listed other paragraphs from that pronouncement they were considered to be “Needs more work.” The assumption was that if the students were writing and doing a full analysis they would have the opportunity to recognize they needed more to properly address the issue. However, if the student did not list the appropriate pronouncement then they were considered to be “Off the Mark”.

The following pronouncements and paragraphs were deemed to be sufficient to have a “Correct’ response.

#### Question 1

The question refers to a situation that the attorneys define as “probably liable” and an estimated range is provided. Then it asks for an opinion based upon GAAP for the proper accounting and disclosures.

FASB 5 paragraph 8:

An estimated loss from a loss contingency (as defined in paragraph 1) shall be accrued by a charge to income if *both* of the following conditions are met:

- a. Information available prior to issuance of the financial statements indicates that it is probable that an asset has been impaired or a liability has been incurred at the date of the financial statements. It is implicit in this condition that it must be probable that one or more future event will occur confirming the fact of the loss.
- b.
- c. The amount of the loss can be reasonably estimated.
- d.

FASB 5 paragraph 9:

Disclosure of the nature of an accrual made pursuant to the provision of paragraph 8, and in some circumstances the amount accrued may be necessary for the financial statements not be misleading.

FASB 5 paragraph 10:

If no accrual is made for a loss contingency because one or both of the conditions in paragraph 8 are not met, or if an exposure to loss exists in excess of the amount accrued pursuant to the provisions of paragraph 8, disclosure of the contingency shall be made when there is at least a reasonable possibility that a loss or an additional loss may have been incurred. The disclosure shall indicate the nature of the contingency and shall give an estimate of the possible loss or range of loss or state that such an estimate cannot be made.

The underlined words are emphasis added for this article. These words explain the accounting, no accrual because a loss cannot be reasonably estimated just a range, and the required disclosure.

## Question 2

The question refers to bonds maturing in 10 years which have been sold at a discount. The CFO wants to amortize the discount using the straight line method.

APB 21 Paragraph 15:

*Amortization of discount and premium.* With respect to a note which by the provisions Opinion requires the imputation of interest, the difference between the present value and the face amount should be treated as discount or premium and amortized as interest expense or income over the life of the note in such a way as to result in a constant rate of interest when applied to the amount outstanding at the beginning of any given period. This is the "interest" method described in and supported by paragraphs 16 & 17 of APB Opinion No. 12....

APB 12 paragraph 16:

The objective of the interest method us arrive at a periodic interest cost (including amortization) which will represent a level effective rate on the sum of the face amount of the debt and (plus or minus) the unamortized premium or discount and expense at eh beginning of each period. The difference between the periodic interest cost so calculated and the nominal interest on the outstanding debt is the amount of periodic amortization.

Both paragraphs support the concept of an effective rate of interest being level over the term of the note rather than a level amount of amortization which would be determined using the straight line method.

bond	attraction	covenant
bonds	link	pledge
bonded	union	contract
bonding	surety	chains
link	collateral	constraint
binder	security	restriction
connection	obligation	join
tie	promissory note	bind
relation	guarantee	

<b>Table 2 - Contingencies Clusters</b>		
<u>Terms</u>	<u>Correct SFAS or APBO</u>	<u>Correct Paragraph(s)</u>
<b>Cluster based on 50 dimensional matrix</b>		
litigation	X	
loss contingencies	X	X
prior period adjustments		
contingencies	X	X
<b>Cluster based on term-document matrix</b>		
adjustments related to prior interim periods of the current fiscal year		
change in accounting estimate		
errors	X	
renegotiation		
contingencies	X	X

<b>Table 3 - Interest Rates Clusters</b>		
<u>Terms</u>	<u>Correct SFAS or APBO</u>	<u>Correct Paragraph(s)</u>
<b>Cluster based on 50 dimensional matrix</b>		
banks		
bonds	X	
interest rates	X	X
markets	X	
maturity	X	X
notes	X	
<b>Cluster based on term-document matrix</b>		
bonds	X	
cash flow		
interest rates	X	X
maturity	X	X
principal		

<b>Table 4 - Comparison of Cluster Term Frequencies</b>			
Question 1		Question 2	
<b><u>Term</u></b>	<b><u>Hits in FARS</u></b>	<b><u>Term</u></b>	<b><u>Hits in FARS</u></b>
contingencies	245	interest rates	365
loss contingencies	90	maturity	339
litigation	54	bonds	152
errors	25	markets	160
		notes	502

Question	Correct Response	Response Needs More Work	Response Off the Mark	Total
1	9	17	37	63
1h	19	19	26	64
2	15	7	42	64
2h	18	13	32	63

<b>Table 6: Test Results Summary</b>		
	Hypothesis 1	Hypothesis 2
Q1	Reject ( $p = .0182$ )	Reject ( $p = .0207$ )
Q2	Do not reject ( $p = .2548$ )	Reject ( $p = .0451$ )