

# Scaffolding Tax Professionals Using the Original Tax Code<sup>†</sup>

Roland Hübscher  
Department of Information Design and Corporate Communication  
Bentley College  
175 Forest Street, Waltham, MA 02452-4705, U.S.A.  
rhubscher@bentley.edu

Brigitte Mühlmann  
Department of Financial Planning and Taxation  
Bentley College  
175 Forest Street, Waltham, MA 02452-4705, U.S.A.  
bmuehlmann@bentley.edu

Mamiko Ono<sup>‡</sup>  
Monitor Group  
Two Canal Park, Cambridge, MA 02141, U.S.A.  
mono@marketspaceglobal.com

**Keywords:** Internal Revenue Code, intelligent user interface, problem solving skills, experimental study

## Abstract

The original tax code is the principal source of law. However, because of its enormous size and complexity, most tax professionals rely on secondary sources, which are not accepted as authoritative sources by the Internal Revenue Service (IRS) or the courts. To remedy this situation, either the original tax code (IRC) needs to be vastly simplified or a possibly intelligent user interface to the IRC needs to be provided. Opting for the latter approach, we did a qualitative experiment studying the difficulties tax professionals have using the IRC. The results indicate that the problem is not just the code, but that the tax professionals also lack the necessary problem solving skills. Thus, we suggest that if intelligent support systems should have a long-lasting impact, they should not just help solving these complex problems, they should also support the tax professionals in acquiring and refining the necessary problem solving expertise.

---

<sup>†</sup> Presented at the *Workshop on Artificial Intelligence and Emerging Technologies* at the Annual Meeting of the American Accounting Association in San Francisco, 2005.

<sup>‡</sup> Mamiko Ono participated in the reported research while she was a graduate student in the Human Factors in Information Design program at Bentley College.

## I. INTRODUCTION

The U.S. tax law regulating all aspects related to taxes in the U.S. has grown from 180 pages in 1934 to over 5,000 pages in 2005. Originally, the tax law was meant to be used by everybody, not just by professional tax consultants and the Internal Revenue Service. However, due its complexity, size and representation it is almost impossible for professionals to use the tax code as a day-to-day resource in the form the official version is published. Thus, most tax professionals do not use the original tax code anymore, but rely on secondary sources that are more understandable, e.g., commercial tax services, although their purpose according to (Raabe, Whittenburg, and Sanders 2003) is to direct the researcher to the germane primary sources of the tax law and not to serve as a substitute. Because these secondary sources contain a pre-interpreted version of the code, they cannot be used as a reference in dealing with the IRS or in court.

The well-respected Judge Learned Hand wrote already in 1947 (Hand 1947):

In my own case, the words of such an act as the Income Tax, for example, merely dance before my eyes in a meaningless procession: cross-reference to cross-reference, exception upon exception—couched in abstract terms that offer no handle to seize hold of—leave in my mind only a confused sense of some vitally important but successfully concealed purport, which it is my duty to extract, but which is within my power, if at all, only after the most inordinate expenditure of time...

Eustice observed that the code then was “almost childlike in its simplicity when compared to the ‘modern’ version of our statute” (Eustice 1989). This increase in volume and complexity has serious implications. According to Selmrod and Bakija (2004) “individual taxpayers spend as much as 3 billion hours of their own time on tax matters, or about 27 hours per taxpayer on average. That is the equivalent of over one-and-a-half million full-time (but hidden and unpaid) IRS employees!”

Although there is regularly talk about simplifying the Internal Revenue Code, there is little hope that, no matter what changes, the IRC will indeed become usable in the near future. Thus, to make the original code usable for tax professionals, we need a user interface to the IRC that assists the tax professional intelligently in the effective use and application of the IRC. For such an intelligent user interface to have a broad impact, it needs to support performance, namely solving tax problems based on the IRC, and education, namely learning to understand, use and apply the tax code effectively.

For instance, knowledge of the structure of the IRC improves research performance under certain conditions (Roberts and Ashton 2003), and code section knowledge, which allows direct access to relevant authority, positively affects tax research performance (Barrick 2001). Cloyd (1995; 1997) provides results suggesting that knowledge of relevant tax laws influences the manner in which tax professionals perform information search. He also found that effort can substitute for knowledge in performing informing information search tasks.

Instead of building yet another tax expert system, we insist that first we need to understand what keeps tax professionals from using the IRC effectively. Obviously, the code has become something like the abominable tax monster, but maybe there is more to it. Human problem solving is an area dealing with these kinds of cognitive tasks and the problems associated with them (Sternberg and Frensch 1991).

Therefore, we conducted a qualitative study of how advanced law tax students solve tax cases and compared their approach to how an accomplished tax law expert does it. We discuss the results and the implications, especially under the consideration that an intelligent computation assistant must support performance and education of the tax experts.

### **Use of Original Tax Code**

An informal survey among tax professionals resulted in a unanimous agreement that the original tax code should be used, instead of secondary sources, and that this is practically impossible because it is huge and complex.

Interpreting the IRC is difficult due to the code's complexity (Raabe, Whittenburg, and Sanders 2003): Several of the sentences in the code are over 300 words long, yet a single phrase may have unexpected outcomes for the taxpayer. A researcher, in his or her initial review, may find the topical index, which is included by most publishers of the code, a useful tool in locating a starting point or the relevant code section. In reading, interpreting, and evaluating a selected code section, the tax researcher must be especially critical of the language used throughout the section. Many, if not most, code sections contain a general rule, followed by specific conditions that must be satisfied in order to apply the provision, and situations under which the taxpayer is excepted from the general rule. In some cases, the exceptions to the general rule are further modified to provide for exceptions to the general exceptions. Moreover, some exceptions to a code section are addressed not within the same section, but in another section of the code. Therefore, all relevant provisions must be read carefully.

In addition to being aware of the required conditions for application of a section, as well as the exceptions thereto, the researcher must be aware of the definitions of terms used in the section; pertinent definitions may be given within the section or in some other provision of the code. These definitions may be significantly different from the common use of the term. In all cases, the researcher should avoid jumping to premature conclusions until a thorough analysis of all relevant code sections has been completed.

The researcher should be aware of the fact that the language of the IRC is not made out of one piece like a Michelangelo statue, but is the result of a certain trial-and-error process. Thus, when analyzing a provision that recently has been changed by Congress, a researcher must be very careful to cross-reference all of the uses of terms whose definitions have been affected by the new law. Often, Congress does not use the care necessary to ascertain that all of the "loose ends" of the new provisions have been tied up. In recent years, almost every major change in the tax law has been followed by a "technical corrections act" to remove errors in implementing and interpreting the new provisions of the law, as well as to clarify problems that arise in integrating the new provisions with the existing provisions of the code. These technical corrections add to the number of updates of the IRC which typically take place at least once a year.

Because the provisions of the IRC change frequently, the researcher must be aware of the effective dates of the various changes to the law. A provision may not go into effect immediately upon its adoption by Congress. The date of the act with which the change in law is passed is not always indicative of the effective date of the provision. Often, various provisions under the same tax law will become effective on different dates and in fact, may have effective dates that precede the date of the tax act.

Similarly, when a provision of the tax law is deleted from the code, the provision may be left in effect for a designated period of time before it actually expires. Transitional rules may also apply. The effective date for a change in the tax law usually may be found in the explanation of

the Public Laws, which follows the pertinent code section. In some cases, the researcher may need to look to the explanation under another code section for the effective date of a provision.

Finally, the tax researcher must be aware that not *all* of the answers to a tax question will be found in the code. The code may be silent concerning the problem at hand, the application of code language to the fact situation at hand may not be clear, or code sections may appear to be in conflict. Thus, the researcher must look for an answer from other sources, such as tax treaties, judicial decisions or administrative sources of the law. Alternatively, the controlling law may be found in other parts of the code, such as tariff or bankruptcy laws. Nonetheless, as a general rule, the IRC as the principal source of tax law should be consulted first.

We also have found that many tax professionals add notes and references into their code and that annotated code is really the proper source for the tax professional since the code does not exist in a vacuum.

### **Purpose and Impact of Intelligent Tax Systems**

There seems to be a general agreement that intelligent systems, specifically expert systems, in law have two purposes, one related to performance and the other to education. The system needs to make knowledge and expertise widely available so that, first, the correct results can be found in a short time even by professionals who don't have in-depth understanding of the specific issues and second, to let those practitioners learn from interacting with the systems.

According to the literature, the expert systems in the tax domain have indeed proven their value on the performance side. However, there is no evidence for educational advantages. Although some people suggest they are indeed educationally beneficial (Susskind 1987), the arguments are weak and everything in the area of education technology indicates that the systems can be expected to have, at best, a minimal positive effect. This is especially true if systems are supposed to replace and not just assist tax experts as Fisher (1997) proposes.

Systems like THUMPER (Swaffield 1991), TAXADVISOR (Michaelsen 1982), TAXMAN (Michaelsen and Messier 1987) and ExperTAX (Newquist 1987) generally ask for input parameters and then will provide an answer. An explanation facility sometimes provides information about how the reasoning engine has come up with the various reasoning steps and why it requests certain input data. Although these systems have been reported as being successful with respect to performance, this passive "participation" of the user cannot result in any interesting learning (Guzdial et al. 1996; Puntambekar and Hübscher 2005).

For the learner to learn from the interaction with a system, the system must make the learner an active participant who also makes decisions. Thus, the learner is actively involved in decision making and the acquisition of the necessary understanding of the tax code, its meaning and implication and the required data. Over time, the learner will be able to take on more and more ownership of the problem solving process and the system changes its role from the master in a master-apprenticeship relationship to an assistant.

If these systems are to have also an educational impact, then they need to be designed from the very beginning on with this goal in mind. Learning unfortunately doesn't just happen as a nice side effect. This requires to understand what the specific problems are that learners have with the material, in our case, with the original tax code.

Thus, we are interested how a tax professional needs to be scaffolded (Puntambekar and Hübscher 2005), i.e., supported solving the tax problems so that the professional keeps learning about the tax law and the methods to solve these kinds of problems. Just providing a system that solves the tax problem will possibly lead to a correct answer to the tax problem but will result in

minor learning at best. Furthermore, the expert does not learn to recognize unreasonable answers given sometimes by the system. Imagine we would give our second-grade students calculators and that's all they had to use from then on. Although the students were able to come up with the correct result of some non-trivial addition or multiplication, they could neither understand whether the result makes sense nor would they be prepared to ever learn even just slightly higher-level mathematics.

### **Overview**

First, we present the experiment to study how tax professionals solve tax problems. We discuss the results and suggest how they should influence the design of systems that are expected to have an impact on performance and education in tax problem solving.

## **II. STUDY**

The goal of the study was to understand how tax experts solve tax problems and what difficulties they have using the original tax code. Although this study used a simplified situation, the outcome was quite eye opening.

### **Method**

Graduate tax law students were recommended by a professor not associated with this project from one of his classes. He was aware of the general topic “problem solving using the original tax code” of the study, but not with any of its details. Of the four recommended students, three participated. The students did not receive any monetary compensation. This seemingly small number is adequate for a qualitative study where the data is studied in minute detail. These students and one tax law professor as an initial baseline participated in the study. The students' experience in tax law ranged from being a tax consultant for several years to “just” being a graduate tax law student with experience in tax-unrelated consulting and journalism, respectively.

The subjects were given the relevant six pages of IRC's code sections one week before they had to solve a tax problem. This allowed them to familiarize themselves with the code sections without knowing the actual problem. The problem was selected to be relatively simple, without being too trivial. The problem statement read as follows:

Carl is a research assistant and Ph.D. student at MIT. He receives \$10,000 salary for 1,000 hours of service over a nine-month period. In addition, Carl receives a waiver of \$15,000 for tuition. The tuition waiver is available to all full- and part-time employees. The fair market value of Carl's services is \$24,000. How much is Carl required to include in his gross income?

The subjects had thirty minutes to find the answer. They were allowed to use any tool they wanted to bring with them including calculators, annotated tax code, paper and pencil, etc. We videotaped them in the usability laboratory and asked them to think aloud (Nielsen, Clemmensen, and Yssing 2002) while they were solving the problem, i.e., they were asked to describe what they were doing, searching for, thinking about, etc. When a subject would become quiet for too long a time, the experimenter encouraged the subject to keep talking.

## Results

After the videotapes were transcribed and coded, all three authors separately and then as a group analyzed the data. The goal was to understand the problem solving strategies employed by the subjects and potential difficulties they had while solving the problem.

Originally, we were worried that the problem would be too easy for the students and that we might not find much evidence for the difficulty of solving a tax problem. Graduate law students are expected to be good at reasoning and making concise arguments. Furthermore, they only had to deal with a minute part of the overall tax law making it much easier to find the relevant information. However, although all subjects arrived at the same correct answer, studying their activities showed that most of them had huge problems, whether they were aware of them or not.

The tax expert had a clear problem solving strategy: (1) Skim the problem statement generating a list (in the head without writing anything down) of facts, assumptions, and subgoals; (2) solve the subgoals in an order to minimize the effort, i.e., those tasks (or subgoals) had a higher priority that could lead to a cutoff in the search tree. The strategy was strongly influenced by the structure of the tax code and the expert's extensive domain knowledge. The tax code was read in an efficient expectation-based manner, i.e., based on the current problem solving state, the expert was looking for specific information based on the knowledge of how the tax law is generally written and structured at the section level. The expert did neither question the decisions made during the problem solving process nor the solution, but made some comments reflecting on the progress, e.g., by explicitly marking the start and accomplishment of a subgoal.

The problem solving behavior was much different for the other three subjects, the students. They all followed a data-driven approach, i.e., they started with the data from the problem statement, and then looked for a paragraph in the tax code that might apply. To find an applicable paragraph, one subject was jumping around in the code quite randomly, another followed a linear strategy reading through the tax code applying the next matching paragraph, and the third subject had the few pages of tax code memorized and thus found a matching paragraph very quickly. However, all four subjects had problems with the fact that they could only focus on one part of the code at a time which is problematic because there are so many references forcing the users to go back-and-forth between several parts of the code.

Two students made several comments implying that they did not really trust their own approach or the result, while the third did not make any meta-comments at all and seemed to just follow (blindly?) a procedure. In general, the three less experienced subjects double checked their results by skimming through most of the tax code to see whether there might be a paragraph that applies to the problem but they had not used.

The three main differences between the expert and the less experienced users (graduate tax law students, nonetheless, some with extensive tax experience) were:

- The expert's strategy was much more driven by the tax code's structure and the domain knowledge, whereas the other subjects used a more data-driven approach.
- The expert used an expectation-driven skimming/reading of the tax law, whereas the other subjects used highly inefficient strategies (random, linear, memorized).
- The expert was actively reflecting on the problem solving progress, whereas the other subjects either did not reflect at all or were full of self-doubts.

## Discussion

Apparently, the subjects with less expertise were not able to apply an effective reasoning strategy to solve the tax problem. Instead of using the structure of the tax code at the level of sections and expectations based on the understanding of the code, they used an inefficient data-driven approach. This is a problem considering the size of the code and the fact that tax consultants are normally paid by the hour. The fact that all subjects came up with the correct answer does by no means justify their approach. They were lucky and having to rely on luck when taking care of tax issues is playing with fire.

This less efficient approach is also more error prone. Incorrect parts of the code are applied and too much time is spent on trying to answer irrelevant questions. This happens because subjects were not able to cut off large chunks of the search tree. For instance, instead of spending time on computing how much can be deducted from the gross income, it would have been more efficient to realize that this deduction does not apply since certain conditions were not met.

The study suggests that the following issues are serious problems for using the IRC.

- The sheer size, and growing, of the tax code makes it difficult to find the relevant information. It also makes it difficult for the problem solver to know that all the relevant parts have been considered.
- Structure that can be used to guide the problem solving process can only be found at the low level of sections. At higher levels (titles, subtitles, chapters, subchapters and parts) that support is minimal to non-existent.
- Without extensive domain knowledge and tax-solving experience a problem solver is missing the necessary heuristics to reduce the size of the vast search space.
- The representation of the tax code does not support the highly non-linear reading process that is required during problem solving (see also Judge Learned Hand's and James Eustice's quotations in the introduction).
- Even advanced graduate law students have difficulties to solve problems that require non-trivial problem solving strategies.

How relevant are these results given that the experiment used graduate students as subjects and they only used a fraction of the IRC? The case could be made that more experienced tax professionals become better problem solvers on the job. Although this may be the case, and further studies are planned with these kinds of subjects, people tend to stick to the method they originally learned and are familiar with, except if they are explicitly told to change their approach. That will probably rarely happen since nobody is aware that their problem solving strategy is wrong. Note that "inefficient" is to tame a characterization since the methods applied by the subject may result in incorrect results if one realistically assumes that they have certain time restrictions when solving a tax problem. Using just a fraction of the code did probably result in improved performance and the study planned with "real" tax professionals will use a larger part of the code or possibly even the whole IRC.

## III. CONCLUSIONS

The experiment has uncovered some serious problems that graduate tax law students have. We suggest a two-pronged approach to remedy this situation. First, the education of these

tax experts must take the students' problems into consideration and make sure that the graduates will have the appropriate skills. And second, intelligent systems supporting tax experts need to not just concentrate on performance results but also on their educational impact.

It would have important practical, not just academic, advantages if tax experts would use the IRC instead of relying on secondary sources. However, in its current form this is not feasible. The IRC is an incarnation of the user-centered design's antithesis.

Instead of waiting for changes simplifying the code, an approach prone to be more successful is to design an intelligent user interface to the IRC that should not only take into consideration the code itself, but also the cognitive characteristics of the users and the two dimensions along which success will be measured: performance and education.

An intelligent interface can be seen as a powerful tool for the user to solve a difficult problem, or it can be viewed as a problem-transforming tool that allows the user to solve a simpler problem than the original one. The existing tools in the tax law domain, mostly expert systems, are of the former kind. If the tool should also support the tax expert in becoming more knowledgeable and a better problem solver, then we should not be interested in the solution alone, but the user should be involved in the decisions, albeit with the assistance from the tool. The goal of this approach is to provide scaffolding-like support (Puntambekar and Hübscher 2005) where the user, over time, learns to use the tax code better with less and less help of the system resulting in a more efficient and competent tax problem solver who will need the crutches provided by the system less and less.

## BIBLIOGRAPHY

- Barrick, John A. 2001. The Effect of Code Section Knowledge on Tax-Research Performance. *The Journal of the American Taxation Association* 23:20.
- Cloyd, C. Bryan. 1995. Prior Knowledge, Information Search Behaviors, and Performance in Tax Research Tasks. *The Journal of the American Taxation Association* 17:82.
- . 1997. Performance in Tax Research Tasks: The Joint Effects of Knowledge and Accountability. *The Accounting Review* 72:129.
- Eustice, James S. 1989. Tax Complexity and the Tax Practitioner. *Tax Law Review* 45:7.
- Fisher, Rodney. 1997. Determination of residence status for taxation law: development of a rule-based expert system. In *Proceedings of the 6th International Conference on Artificial Intelligence and Law*. Melbourne, Australia: ACM Press.
- Guzdial, Mark, Janet L. Kolodner, Cindy E. Hmelo, N. Hari Narayanan, David Carlson, Noel Rappin, Roland Hübscher, Jennifer Turns, and Wendy C. Newstetter. 1996. Computer Support for Learning through Complex Problem Solving. *Communications of the ACM* 39 (4):43-45.
- Hand, Learned. 1947. Thomas Walter Swan. *Yale Law Journal* 57:167.
- Michaelsen, Robert H. 1982. Application of an expert computer system to federal tax planning. *ACM SIGART Bulletin* (82):18-19.
- Michaelsen, Robert H., and William F. Messier. 1987. Expert Systems in Taxation. *The Journal of the American Taxation Association*.
- Newquist, Harvey, III. 1987. The Making of Tax Expert. *AI Expert* 1 (2):149-167.
- Nielsen, Janni, Torkil Clemmensen, and Carsten Yssing. 2002. Getting access to what goes on in people's heads: Reflections on the think-aloud technique. In *Second Nordic Conference on Human-Computer Interaction*.

- Puntambekar, Sadhana, and Roland Hübscher. 2005. Tools for Scaffolding Students in a Complex Learning Environment: What Have We Gained and What Have We Missed? *Educational Psychologist* 40 (1):1-12.
- Raabe, William A., Gerald E. Whittenburg, and Debra L. Sanders. 2003. *West's Federal Tax Research*. 6th ed: South-Western College/West.
- Roberts, Michael L., and Robert H. Ashton. 2003. Using Declarative Knowledge to Improve Information Search Performance. *The Journal of the American Taxation Association* 25:21.
- Selmerod, Joel, and Jon Bakija. 2004. *Taxing Ourselves, A Citizen's Guide to the Debate over Taxes*. Third ed: The MIT Press.
- Sternberg, Robert J., and P. A. Frensch, eds. 1991. *Complex problem solving: Principles and mechanisms*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Susskind, Richard E. 1987. Expert systems in law: out of the research laboratory and into the marketplace. In *Proceedings of the 1st International Conference on Artificial Intelligence and Law*. Boston, MA: ACM Press.
- Swaffield, Gail. 1991. THUMPER: an expert system for stamp duty. In *Proceedings of the 3rd International Conference on Artificial Intelligence and Law*. Oxford, England: ACM Press.