The University of Sydney

Sydney Law School

Legal Studies Research Paper
No. 06/41

August 2008

Testing Alternative Legal Paradigms: An Experiment in Designing Tax Legislation

Graeme Cooper & Michael Wenzel

Testing Alternative Legal Paradigms:

An Experiment in Designing Tax Legislation

Graeme S. Cooper and Michael Wenzel

CONTACT AUTHOR:
Graeme S. Cooper
School of Law
University of Sydney
175 Phillip Street
Sydney NSW 2000
Australia
Phone: 61 2 9351 0310
Fax: 61 2 9351 0200
E-mail: G.Cooper@usyd.edu.au

Graeme S. Cooper is Professor of Taxation Law in the Faculty of Law at

the University of Sydney, Australia.

Michael Wenzel is a Senior Lecturer at the School of Psychology,

Flinders University, Adelaide, Australia. At the time this research was conducted,

he was a Fellow at the Centre for Tax System Integrity, Research School of Social

Sciences, Australian National University, Canberra, Australia.

ABSTRACT

This article reports on empirical research undertaken to test the claim made in a law reform project that citizens could be made more certain of their legal obligations by changing the legal paradigm used to express their rights and obligations. Our research tested a number of hypotheses involving different formulations of the claim being made. We find that the alternative paradigm being presented was inferior to current practice and offer some reasons which would explain our results and the significance of this work for other areas of legal research.

INTRODUCTION

Consider this proposition: every year, financial institutions provide vast amounts of capital to customers. They do so using relatively well-known legal paradigms: unsecured loans, mortgages, debentures, hire-purchase arrangements, finance leases, corporate bonds, bills of exchange, promissory notes, and so on. And yet, every year, these institutions and their customers have disputes about the meaning and extent of their obligations. On the one hand, lenders find they cannot recover some of their funds, not because the borrower is insolvent or the security proves inadequate, but because the terms of the arrangement are found by a court to mean something other than the lender believed. On the other hand, borrowers sometimes find that their liabilities are greater than they had believed--the deal struck does not reflect their understanding of it.

Consider now a second proposition: modern corporate finance tells us that most financial positions can be replicated using derivative instruments--ownership can be replicated using a call option, an immediate sale can be replicated using a put and call option, and so on (Brealey, Myers, and Allen 2006; Smith 1990). In short, many arrangements that are currently constructed one way could be constructed in other ways.

Now the question: given that many legal arrangements might be constructed from a range of different legal paradigms, might the effectiveness of an arrangement be improved by using one paradigm rather than another? For

example, would borrowers and lenders be better served by abandoning some of

the traditional legal instruments, and the legal paradigms they reflect, and

substituting another instrument based on a different legal paradigm? Rather than

lending money to a home buyer and taking security for repayment in the form of a

mortgage--a power to seize and sell the security--would the bank or customer or

both, be better served if the home buyer were formally to sell the house to the

bank and take a call option to repurchase it at a designated price, enforceable so

long as the customer continued to pay periodic option fees? Would the parties

understand their obligations more thoroughly if the deal were constructed and

presented to them in this manner? Would there be fewer disputes between banks

and borrowers? Would the remaining disputes be cheaper to resolve? Would

third parties understand the parties' respective positions more fully if the deal was

presented to the outside world in this new form?

Other private law arrangements would lend themselves to the same

question: would a casualty insurance policy be more effective if constructed as a

series of yearly put options held by the insured? The same question could also be

asked in a public law context. If the government were proposing to regulate

lending practices, would borrowers and lenders be better served if the government

constructed the regulatory regime in a way which replicated the paradigm of a

loan with security, or adopted the paradigm of a sale with an option? Would

corporations' law be better comprehended if the government abandoned the

traditional idea of limited liability and expressed the same outcome through the paradigm of a put option held by shareholders over corporate assets? Would an environmental law regime be more effective if it simply expressed a range of prices for emissions of various kinds, or imposed a series of excise taxes on the production of various pollutants (Posner 2003)?

Of course, substituting a new paradigm for an established one may solve one particular set of issues, but it might also create problems that did not previously exist, both during transition and in the long run. To continue our mortgage example, even if some rights of the parties *inter se* were slightly clearer, other questions would undoubtedly arise. Who would now be responsible to insure the property? Under the loan with power to seize paradigm, it seems plausible to suspect that the borrower, as owner, would assume that this responsibility fell to her. But if the bank is now the owner and the borrower has merely occupancy rights with a call option, would the parties believe that the bank should now insure the property? During transition to a new paradigm, the parties may not think about it. If they did consider it and concluded that the responsibility would initially lie with the bank but that this was not the intended outcome, presumably the responsibility could be shifted back to the borrower by contract, but the process would now be slightly more cumbersome. The loan paradigm, whether by implication or established usage, conveyed the correct message that the borrower had the responsibility to insure the property; if the

option paradigm conveyed the reverse and would have to be negated by contract, an additional complication has been created and a new opportunity for misunderstanding and error has crept in.

The same issue would arise with respect to dealings involving third parties. For example, who would now be responsible for the payment of property taxes? This time, the initial allocation of the obligation would be set by the terms of the property tax statute. If the property tax statute made the "owner" liable to pay property tax assessed on the property, the taxing authority will need to know whether it is the bank or the occupant who meets that description. Once that is resolved, there would need to be a consequential adjustment between the parties if the cost lie with the bank but this was not intended. If the parties concluded that paying the property tax was a cost of the borrower, presumably a liability to pay an amount equal to the property tax could be added to the borrower's debt by contract, but the process is now more cumbersome. It is not hard to think of other instances where similar confusion might arise: Who has liability to compensate third parties injured while on the property; do the rights of the occupant pass, on death, to the inheritors of their real property or the inheritors of their personal property? Whatever the set of issues that the change in paradigm was meant to solve, it is plausible to suspect that another set of issues would also be created. The possibility that there would need to be more pervasive and permanent steps to achieve the same outcome is a real concern. While we might easily accept that

many legal arrangements could be constructed using one of several different legal paradigms, it must be an open question whether the effectiveness would be improved by shifting from one paradigm to another. The topic of this article is that question.

The immediate impetus for this article was a recommendation of the Australian Review of Business Taxation (RBT) (Review of Business Taxation 1999a) to reconceptualize the base of the Australian income tax. The RBT recommended replacing the current expression of the tax base with a formulation that it called the "tax value method" (TVM). Tax value method was not simply a new phrase--it was a new paradigm; it constructed the tax base using different legal constructs. And yet, the RBT assured us, the new legal paradigm would lead to the same outcomes, with greater certainty and at less cost. The tax system would be more "coherent and durable," would display less "uncertainty and complexity," would be more "consistent, transparent and sustainable," be more "enduring and robust," and be more flexible (156).

The claim that TVM would lead to the same outcomes, with greater certainty and at less cost, seemed doubtful to us. None of the assertions were justified by argument, theory, or data in the RBT's report; it was asserted as self-evident. Yet it seemed to us that both the RBT's confidence in its propositions and our skepticism about some of the claims were testable – it should be possible to conclude whether this was indeed a better mousetrap, or just another

6

mousetrap. This article reports how we tested the claim of greater certainty, what we discovered, what reasons might explain the results, and what those reasons might mean for law and legal systems more generally.

Before we began our project, we had expected to discover similar projects where the same question was posed: given a choice between legal paradigms, does one of the alternatives result in a greater degree of certainty in its interpretation, either because it is inherently less complex, creates fewer questions of interpretation, or is intuitively easier to comprehend and administer? This question must have been pondered by others; it must surely crop up almost everywhere in public law and private law.

It came as something of a surprise to find that lawyers and social scientists appear not to have asked our question. There is a large experimental literature--in the field of public policy (Dunn 1998) and in public economics (Kling 2007; Alm and Jacobson 2007)--but understandably it does not focus on what is essentially a legal question. Even in the legal and sociolegal literature, there is little about how and why to choose one legal paradigm over another when several are available; the growing empirical and experimental literature in law does not yet contain a stream that focuses on the way that transactions or regulatory regimes are constructed (Genn 2006; *University of Illinois Law Review* 2002). One field where we expected to see some exploration of similar questions is in the commercial lawyering literature, but its focus tends to be on other questions--

principally reducing transaction costs--reducing the scope for disputes through choice of language, avoiding unwanted regulatory burdens, the opportunities for selecting the range of legal regimes that might apply to a transaction, and so on (Gilson 1987; Fleischer 2002; Margolis 1998). The closest cognate literature in law involves the analysis of language. We examined the literature involving jury instructions. A substantial amount of research examines how juries or mock juries interpret different sets of instructions, or, more commonly, fail to properly understand a single set of instructions (Cronan 2002; Garvey, Johnson, and Marcus 2000; May 1995; Power 1999; Steele and Thornburg 1988). Some of these studies include discussions of how the organization, rather than the syntax of a particular instruction set, makes the instruction easier to comprehend (Steele and Thornburg 1988; Diamond and Levi 1996; Charrow and Charrow 1979; see Elwork, Sales, and Alfini 1982). Another obvious analogue is the work on plain legal language, which evaluates attempts to increase the certainty and reduce error costs in legal drafting (Mellinkoff 1983). Ultimately, we considered this literature offered us no assistance because the focus of the literature is on a different question. It offers guidance on how to craft documents that are easier to understand, and it frequently reports experiments and empirical evidence to support these claims. Our question is different; put simply, we are interested in knowing whether a call option works better than a mortgage, an insurance policy

works better than a put option, a tax regime works better than a catalogue of environmental prohibitions, be they well or poorly written.

EXPRESSING THE INCOME TAX BASE

In this section we briefly describe the paradigm underlying how Australia's income tax law is currently constructed.

The basic design elements of Australia's income tax follow the pattern of every income tax. Taxpayers must pay tax annually at stipulated rates on their taxable income (Income Tax Assessment Act 1997, section 4-1; section 4-10). The tax base--taxable income--is a net figure, the amount remaining after subtracting deductions from assessable income (section 4-15). At its simplest, the tax base will represent the excess of various amounts received by the taxpayer during the year over various amounts spent by the taxpayer during the year. Complexity arises because, for a variety of reasons, not all receipts will be income, not all payments will be deductions, and not all receipts or payments will be allocated to the years in which they are made.

The rules that define income, deductions, and timing are always difficult, but the claims for TVM revolved around reconceptualizing these three elements. Looking first at the assessable income element of the tax base, in Australian tax law, there is general agreement that income means three broad classes of amounts: amounts received as a reward for performing services as an employee (wages, fringe benefits, performance bonuses, and similar employee benefits);

profits from carrying on a business; and amounts received as a return from investments (dividends, interest, rent, royalties, and so on). Some common omissions are gifts and other windfalls, inheritances received, prizes, gambling winnings, the proceeds of crime, damages and compensation for loss or injury, the principal of a loan, and gains made on the redemption of liabilities, although there are exceptions to each of these examples.

In Australian tax jurisprudence one important omission from the notion of income is a gain made on the disposal of a capital asset, although this class of transactions has been subjected to tax under a special statutory regime for taxing capital gains since 1985 (Div. 100 ff.). While the profits made on the sale of capital assets may not be ordinary income, they are nevertheless taxed.

With respect of the deduction element of the tax base, the tax statute provides that an outlay that is relevant to earning income, and which does not procure an enduring asset, is deductible (section 8-1). The idea that the outlay is incurred in earning income means that consumption expenses will be excluded from the class of outlays that reduce the tax base, although defining personal consumption is notoriously difficult (Halperin 1974; Klein 1966; Samansky 1981). Disputes typically arise about the cost of clothing, commuting, child-care while parents work, business meals, home-office expenses, travel and accommodation while on business trips, personal grooming, medical expenses, self-education expenses, and so on.

Finally, there are the timing rules. These timing rules typically exist in two places. First, there is the distinction between cash and accrual accounting for the recognition of income and expenses. Cash accounting will record a receipt or payment in the period in which the receipt occurs or the payment is made. Accrual accounting instead records income and expenses sometimes before and sometimes after receipt or payment. A second set of timing rules is needed to shift the money spent to acquire assets that contribute to generating income in later periods. Inventory rules will typically defer the deduction for the cost of the inventory until it is sold. Depreciation rules will spread the cost of equipment and machinery over its useful life. The cost of nondepreciating capital assets will be recognized when sold or destroyed. Each of these regimes has the effect of deferring recognition of the cost from the period in which the outlay was made to one or more later periods.

THE TVM PROPOSAL

The system described above is undoubtedly arcane and more than a little obtuse. It has a certain logic, but no one would have designed such a system had she the luxury of starting with a clean slate. Instead, it is the inglorious result of the usual pressures to which any tax system is subject over time--legislative grafting, historical legacies, taxpayer creativity, administrative pragmatism, political expediency, judicial error, and so on.

The TVM proposal was first exposed in the 1999 report of Australia's RBT (Review of Business Taxation 1999a, 155) and the draft legislation which accompanied it (Review of Business Taxation 1999b, 1999c). The government announced its "in-principle support" for the TVM proposal and referred it to a working party for further consultation and development with the expectation that the proposal would be implemented by July 1, 2001 (Australian Government 1999). In August 2000, the government announced that it would establish a Board of Taxation and referred the TVM proposal to the board for further development and evaluation (Australian Government 2000). The board eventually recommended to the government, and the government accepted, that the TVM project be terminated. During the extensive consultation on the proposal, the Board of Taxation elaborated the original proposal releasing other public information documents and preparing various drafts (prototypes) of an income tax law constructed in accordance with the TVM paradigm (Australian Government 2002a).[1] The following discussion of various legislative provisions is based on the text of the fourth prototype Bill (Prototype 4) prepared for the Board of Taxation (Australian Government 2002b).

According to the original RBT report, TVM would be superior because of the new paradigm used to express the tax base. TVM would be based on four concepts: receipts, payments, assets, and liabilities.

Prototype 4 constructed the income tax base as "net income" plus "taxable income adjustments" less unused tax losses (Prototype 4, section 6-15). "Net income" was defined as receipts reduced by payments and the effect of the net change to the tax value of assets and liabilities and tax losses (section 6-55). Putting together the various elements, the tax base thus became:

(a) cash received,

(b) *less* cash payments made,

(c) *plus / minus* changes to the tax value of assets during the year,

(d) *plus / minus* changes to the tax value of liabilities during the year,

(e) *plus / less* taxable income adjustments,

(f) *less* tax losses.

The legislation excluded from the calculation any amounts paid or received and any asset or liability that is "private or domestic in nature" (section 222). In other words, no entry at items (a) to (d) should include an amount that was of a private or domestic nature.

Several of these terms received further definition. An asset and liability were both defined, and the "tax values" of individual assets and liabilities were stipulated, in a series of tables which replicated current law. Thus assets that were inventory were given a tax value of cost or market value (section 10-40, Item 2); depreciating assets were given a tax value that was the asset's depreciated value at that time (section 10-40, Item 3; section 72-35); nondepreciating assets were for

the most part given a tax value of cost (section 10-40, Item 6; section 78-20--78-100); the tax value of consumable supplies (such as office stationery) was set at zero, and so on (section 10-40, Item 1; section 68-10).[2]

The differences between the TVM paradigm and the paradigm underlying current law may be less than obvious, but consider two differences. First, TVM imposes a requirement to examine the holding of assets and liabilities. To use a financial accounting analogy, one way of conceptualizing the first difference would be to view current law as directed primarily to analyzing elements in the profit and loss account, while TVM directs attention to the balance sheet (Gammie 2001). That is, while current law focuses primarily on receipts and outgoings, TVM focuses on movements in the values recorded in a fictitious balance sheet, the entries in that balance sheet recorded at "tax value." This means that TVM requires taxpayers to recognize, classify, and value all assets and liabilities each year. It also requires taxpayers to understand the process of transformation of cash to asset and asset back to cash. A second point of difference is that TVM deliberately avoids reliance upon any notion of what "income" means in the popular imagination--it did not employ the ideas of wages, interest, or dividends to define the tax base. This agnosticism was deliberate and seen as a virtue of the system.

Embedded within the TVM formula are the tax base rules and timing rules for TVM. Under the TVM paradigm, the tax base consists of nonprivate receipts,

14

reduced by nonprivate payments, modified by movements during the year to the tax value of assets and liabilities. In the simplest cases, a taxpayer makes income if she receives an amount of nonprivate cash and does not suffer a corresponding liability with an equal tax value. On the other side, a deduction arises if she spends nonprivate money and does not acquire an asset with an equal tax value. Where a payment procures an asset, the time at which the outlay is recouped depends on the rules setting the tax value of that class of asset. If the rules set the tax value of the asset at zero, the net effect will be that the amount spent on the asset is deductible; if the rules set the tax value of the asset at its cost, and do not modify that tax value during the period while the asset it held, the net effect will be that the amount spent on the asset is recovered only on the loss or disposal of the asset.

The intended operation of the TVM paradigm can be demonstrated (and contrasted with the current system) by a few examples. Under current law, employees are taxed on wages because wages are considered ordinary income as the result of rendering services; under TVM, wages would be taxable because there is a receipt that is considered to be nonprivate. Under current law, interest is taxable because it is considered ordinary income as the return on an investment; under TVM, interest would be taxable because there is a receipt considered to be nonprivate. Under current law, money spent on groceries would not be deductible because it is not spent in order to earn income; under TVM money spent on

groceries would be a payment and so prima facie deductible, but it would be a payment of a private or domestic nature and so excluded from the payments that deplete the tax base. Under current law, money spent to purchase a block of land for the head office of a business would not be deductible--it would be considered an outgoing of a capital nature. Instead, the cost of the land would be subtracted from the proceeds of sale to determine the amount of any capital gain made in the year of sale. Under TVM, money spent to purchase the block of land would be a nonprivate payment, and so prima facie deductible, but the payment would give rise to an asset with a tax value equal to the payment. These two elements would offset in the year of payment and there would be no net tax effect. In the year of sale, there would be a receipt, but the land would not appear in the taxpayer's closing assets. The difference--the receipt reduced by the opening tax value of the sold asset--would be taxed in the year of sale. Under current law, the principal of a loan borrowed by the taxpayer would not be considered ordinary income; under TVM, the amount borrowed would be a receipt and so prima facie taxable, but the taxpayer would also be subject to a liability with a tax value equal to the amount borrowed. These two elements would offset in the year of borrowing and there would be no tax effect.

In these examples, TVM replicates the treatment that current law achieves for many transactions--be it the treatment of wages, interest, groceries, real estate investments, or borrowed funds--but employs a different paradigm.

16

COMPARING THE TWO PARADIGMS

The given examples show that the two systems can reach the same results in simple cases, but our concern was with the claim to superiority. TVM should be able to manage better; that is, to prescribe and convey the consequences of a transaction with more clarity, and not introduce problems of its own making, unless they too could be easily and intuitively managed. It seemed to us that TVM did create problems that did not previously exist and the rectifications were both complex and sometimes counter-intuitive. We will examine three problems to illustrate the different structural elements of the two regimes and some of the related adjustments

Reversing the Effects of Mere Receipt or Payment

One difficulty that TVM must address is the prima facie position that every receipt is taxable (Prototype 4, section 6-55). TVM includes all cash and noncash receipts, but under current law gifts and other windfalls would not usually be taxable to the recipient. If TVM were to clarify but not change the tax base, this must be reversed. TVM contains two devices to accomplish this. In order for some gifts not to be taxable and others to be taxable, TVM utilizes the concept of the gift that is "private and domestic" (section 222-25(1)(c)). One can easily imagine a court concluding that a gift from a generous relative is private and so excluded from the receipts that are taxed. A court might also conclude that a gift from an employer unmotivated by personal feelings was not "private and

domestic." The same kind of analysis had to be applied to inheritances received, prizes won, and gambling winnings.

Secondly, TVM has to manage the inclusion of other kinds of receipts such as borrowed funds or the subscribed capital of a company. In the case of a full recourse loan, TVM offsets the receipt by recognizing a liability with a tax value equal to the amount borrowed (section 14-80). This treatment works adequately for full-recourse debt. It is less obvious that it easily accommodates non-recourse or limited-recourse debt, or subscribed capital.

If TVM is overinclusive on the income side, it is also overinclusive on the deductions side. Every payment under TVM is a deduction unless it is private, procures an asset that has a positive tax value, or else achieves the discharge of a liability (section 6-55). For example, payments of local income tax, payments of dividends or other profit distributions, fines and penalties, and gifts made to relatives or unregistered charities would all become deductions. Under current Australian tax jurisprudence, these payments would not reduce taxable income.

Again, the device used in TVM was to exclude any payment that was "private and domestic" in nature. It seems clear that this phrase would apply to payments for groceries, holidays, and sports cars--they are "private and domestic" in nature. But it is not clear that these words would capture payments of dividends, income tax, or fines, and hence further adjustments were needed to reverse the effect of mere payment.

Managing the Articulation Between Receipts and Assets, Payments and Liabilities

The next issue arises from a duplication problem. Under one part of the TVM paradigm (lines (a) and (b)), a taxpayer examines receipts and payments, something akin to cash flows during the year. Under another part of the paradigm (lines (c) and (d)), the taxpayer must compare its assets and liabilities at the beginning and end of the year. There is potential for double-counting confusion here.

Consider these two manifestations of the problem. A taxpayer who receives $1,000 has a receipt that will be prima facie taxable as a receipt, but if the taxpayer then deposits the amount in the bank, the taxpayer also has an asset with a tax value at the end of the year that is higher by $1,000. At first glance, the formula would seem to count the same $1,000 twice: once as a receipt and again in the higher closing tax value of an asset. This is clearly not intended, and a way of handling it was proposed: the taxpayer must be treated as having made a payment of $1,000 when she deposited the amount in her bank account. Payment negates the effect of the receipt--we are left with just the asset to be counted. (This has the curious consequence that saved wages are taxed not as receipts but as an asset.) But it seemed to us that this solution to the double counting problem is less than intuitive. A person who spends money on a car probably regards the price as a "payment," but does a person regard a deposit in a bank account as akin

to spending money? We doubted that many people would conceive of themselves as purchasing their bank account from the bank, but unless they do they will not reach the intended result. If they fail to recognize the payment, some astute taxpayers might then (incorrectly) omit the asset, aware that something is awry, but others may be unaware or unable to resolve the apparent dilemma.

A second example of this double-counting problem arises where a taxpayer receives a gift of $1,000. She has a receipt that will be prima facie taxable, but if it is a mere gift, the receipt is not intended to be taxed and would be excluded from the tax base because it is a "private or domestic" receipt. But what happens if the taxpayer then deposits the gift in a bank account? The taxpayer ends the year with a (nonprivate) asset with a positive tax value. So the gift would be taxed, albeit as an asset. This was not intended, and there are several options for handling this. One would be to view the bank account also as "private and domestic," at least in part. Another solution would be to treat the deposit as a nonprivate payment; that is, while the gift received might be private, depositing it into the bank is not. If these steps are followed, the taxpayer would ignore the gift receipt, treat the deposit as a recognized (deductible) payment, and treat the bank account as an asset with a tax value of the amount deposited. The net effect would be nil: the deposit would offset the effect of the having the asset. But there is a danger--the taxpayer might (understandably) treat the deposit of the gift into

the bank account as private or domestic because the gift itself is private or domestic. This would not lead to the correct result.

These difficulties arise from the need to classify (and sometimes even to generate receipts) payments and assets in order to handle the transformation of receipts to assets. The opportunities for overlooking one or more, and for mismatches in the transformation of one to another, gave us pause to wonder. Do these processes ultimately make the TVM paradigm inferior?

The Multiplication of Taxing Points

The last issue we examine arises from the implicit timing rules embedded in TVM.

We mentioned above that under current law, taxpayers who operate on a cash basis derive income when they receive cash or property. For these taxpayers, the making of a demand for payment is not a relevant tax event. Taxpayers who operate under accrual accounting rules derive income when they are entitled to be paid and demand payment. For these taxpayers, the receipt is not a relevant tax event.

We noted above that under TVM, both receipts and assets are relevant to the tax base, but which one determines *when* income is derived? For example, assume a taxpayer signs a contract to perform a service for a client for a price of $20,000, the price being payable on completion of the work which will take place in the next tax year. There are three items disclosed in these simple facts, each of

which has significance in TVM: the contract for the work is an asset, it imposes

obligations and creates entitlements, and there will eventually be a cash receipt.

The interplay of the recognition of each item and the tax value ascribed to it will

determine whether the taxpayer is taxed at the time of making the contract or at

the time of receipt--will the taxpayer pay tax on $20,000 this year, or next?

Under TVM, the answer to this question is not self-evident--it depends on

the interplay of the definitions of "asset," "liability," "tax value," and "cost." At

the end of the year, the taxpayer has not yet received payment, but she has an

asset in the form of the rights under the contract. However, under TVM this

matters only if the asset has a non-zero tax value. Many assets are given a tax

value of their cost. It would seem appropriate to treat the contract as having no

"cost." If that were the case, the existence of the contract as an asset at the end of

the year would be irrelevant--the fact that it had no cost, and thus no tax value,

would remove it from the income of that year. However, the TVM rules proposed

treating undertaking a liability as the giving a "non-cash benefit" to the client

(section 16-15). The effect of this deeming would be to treat the taxpayer as

having paid an amount to the client; the contract would now have a cost, and thus

a positive tax value (section 10-40(1), Item 6). Recognizing this problem, the

drafters proposed a special regime for so-called "routine rights and liabilities" --

rights and liabilities arising under contracts that were incomplete at the end of a

tax year (section 68-45). The rule treated both the asset and liability as having a

tax value of zero (section 68-10(1)(a)). Thus, TVM required a special regime to solve timing issues.

A similar problem of choosing the appropriate taxing point arises for sales on credit--a taxpayer sells goods to a client for $20,000, payable after delivery. So, is the taxpayer to be taxed at the time of making the contract, at the time of delivery, or at the time of payment? The goods that are to be sold would be an asset, the contract of sale would be another asset, and there will also be an eventual receipt. The contract requires the taxpayer to deliver goods, which would appear to be a liability. Again, a special short-term trade credit regime is needed to achieve the intended outcome, which was that the seller would report as income the anticipated receipt of the price at the time of delivery, being the time when the taxpayer has performed all its obligations, and has the right to receive the payment (section 76-15, Item 2).

These examples demonstrate something of the cause for our skepticism about TVM. TVM can certainly accommodate each of the transactions we described, but the question remains whether the rules in TVM make the entire paradigm more difficult and uncertain than its competitor.

THE EXPERIMENT

In this section we describe the project we undertook to test the claim that TVM would lead to greater certainty than current law. The research was

undertaken by conducting experiments involving the application of each paradigm under controlled conditions.

It was necessary to define more carefully what we were testing for, as the RBT might have been using the term "certainty" in several senses. The RBT might have been asserting that TVM would lead to greater intersubjective consensus, greater subjective confidence, and/or higher levels of accuracy. To us, these were the three most plausible definitions of "certainty" and the most susceptible to something approaching rigorous analysis.

Assuming that these were the most likely meanings, we considered that the validity of the claim could be demonstrated in several ways. TVM would be more certain than current law if (a) more respondents agreed on the same answer to tax questions, or the dispersion of their answers was smaller, under TVM than under current law (*objective certainty*); (b) respondents felt more confident in the correctness of their decisions (outcomes and procedures) under TVM than under current law (*subjective certainty*); (c) more people reached the answer that the legislature intended (*substantive accuracy*); or (d) more people accurately applied the processes which the rules prescribe in their efforts to reach the answer that the legislature intended (*procedural accuracy*).

Experimental Issue

In order to test the claim that TVM would increase certainty we formulated and investigated five specific hypotheses.

First, we tested the prediction that TVM would yield greater *objective certainty*; that is, more people would agree on the same answer under TVM than under current law. This hypothesis is formally stated:

> *H1:* *Compared to the current law, TVM will lead to greater consensus on answers to common tax problems.*

Secondly, we tested the prediction that TVM would yield greater *subjective certainty* in two senses: taxpayers would indicate greater confidence in the correctness of their proposed tax position, or greater confidence that they had correctly followed the steps required by law to determine that position:

> *H2:* *Compared to the current law, TVM will lead to greater confidence in the correctness of answers to tax problems.*

> *H3:* *Compared to the current law, TVM will lead to greater confidence that the steps used to determine the answers to tax problems have been followed correctly.*

The hypotheses just described do not include any reference to reaching "the right answer"--either consensus or confidence would demonstrate certainty. This agnosticism on our part was deliberate; it would be an important, perhaps even a sufficient, achievement if TVM were able to induce more people, or lead people more often, to the same answer without also expecting that it be the "the right answer." Nevertheless, we took the RBT's claim of greater certainty as intended also to imply greater likelihood of reaching the legislature's *intended*

25

outcome. Hence we considered *accuracy* should also be a relevant line of inquiry for this project and decided to test two subsidiary claims: that taxpayers would demonstrate greater agreement with (our view of) the intended outcome to tax problems (*substantive accuracy*), and that taxpayers would demonstrate greater accuracy in applying the steps by which the legislature intended the outcomes to be determined (*procedural accuracy*):

> H4:   *Compared to the current law, TVM will lead to greater frequency in reaching the intended answers to tax problems.*

> H5:   *Compared to the current law, TVM will lead to greater frequency in applying the steps intended to be applied in order to determine answers to tax problems.*

In addition to testing these five hypotheses, we also considered three other questions in order to derive more robust results. We were concerned that complicating factors other than TVM--the design of the research or the identity of the participants, for example--might account for our results. So we asked, in so far as differences were found: Were the differences attributable to either a perception that one presentation was more difficult than the other (or the degree of the participants' prior experience with tax and legal matters), and were any differences universal or limited to identifiable subject areas?

Method

To test the predictions of the hypotheses, we selected participants who had few prior experiences with tax problems. This selection was made to try to evaluate the long-term outcome of a move to TVM, rather than a transitional impact. Second, we wanted to exclude the possibility that experience with current law could compromise the results. We considered that University of Melbourne undergraduate students would be a suitable population because they would have little prior experience with tax issues but were sufficiently intellectually adept to process a dense introduction to tax legislation. To reduce further the possibility that the students possessed extensive prior tax experience, we restricted the sample to students who were not enrolled in either a Law or Commerce degree, who were twenty-five years of age or younger, and who had never completed a Business Activity Statement.[3]

The experimental material was pretested on two groups of participants and given to two tax law experts[4] to review for appropriateness and bias. In addition, the Board of Taxation also sent the material to the TVM Legislative Team. Several changes were made in response to the pretesting and comments.

Campus advertisements invited students to participate in a study of people's understanding of legal texts and legislation. Participants were allocated to one of four sessions at which they were given an oral presentation on tax law, two presentations on the current legislation, and two presentations on the proposed TVM legislation. The text at each presentation was read to participants

27

by a recruited presenter (a graduate student in drama) to ensure a double-blind experiment. Participants could also follow each text verbatim in written handouts and received extracts of the tax law as a further reference source.[5]

The choice of what to instruct and the tasks asked of students were influenced by several factors. First, there seemed to us little purpose in testing areas where TVM involved little explicit structural change. Second, areas of high detail were excluded because the effort needed to convey what was meant was not matched by its importance to common transactions. Third, we also excluded some of the peripheral aspects of TVM. In so far as it was possible, our focus was on the principal structural change that TVM proposed to make; that is, the expression of the capital-income dichotomy and timing rules of the new tax base in terms of cash flows, assets, and liabilities. We focused the presentations and materials on three groups of issues:

> Group A: is there a receipt or a payment, how is the double-counting managed for a receipt that becomes an asset, and how is it managed for a payment that diminishes an asset;
>
> Group B: time of recognition of income and expenses; and
>
> Group C: traditional capital/income issues for income and expenses.

While the categories are not perfectly discrete, they allowed us to be more careful in attempting to identify whether there are any *particular* areas where the TVM paradigm might be more or less effective.

Sixty-nine undergraduate students from various faculties at the University of Melbourne participated in the study. Participants were between eighteen- and twenty-eight-years-old (Mdn = 20); thirty were male, thirty-nine were female.[6] Students enlisted in one of four experimental sessions with fifteen to twenty participants in each session. Two of the sessions were randomly assigned to the TVM condition (n = 37), and the other two were assigned to the current law condition (n = 32). Participants at the sessions did not differ in age, $t(67) = -.08$, ns, or in sex composition, $\chi^2(1) = .66$, ns.

The study varied experimentally the tax legislation to which participants were introduced. One group was instructed in TVM, the other group was instructed in the current legislation (CL). This between-subjects factor will be referred to in the following as *legislation*. For supplementary analyses, three groups of tax problems were distinguished as defined above (A, B and C). *Problem type* thus constituted a second factor that was varied within subjects (that is, each subject was confronted with all three problem types).

After the presentation of the relevant tax legislation was completed, participants were asked to complete a questionnaire. The questionnaire first solicited demographic details of age, sex, and area of study. A second set of questions referred to the participants' impressions of the presentations of the tax legislation and rating of the quality and difficulty of the presentation: "The presentation was… [attribute]" (1 = *not at all*, 7 = *very much*). These questions

were asked to ensure that if, despite the review process, the two experimental conditions differed either in perceived quality or difficulty, these ratings could be included in our analyses and their impact controlled statistically. The analysis of these items yielded two constructs of interest here: *Quality of presentation* was measured by the three attributes *clear*, *well structured*, and *articulate* ($\alpha = .80$); *Difficulty of presentation* was measured by two items, namely *complicated* and *demanding* ($\alpha = .66$). Scale score were obtained by averaging across respective items.

A third set of questions asked participants to rate their prior experience, either first-hand or in their family, with tax issues, business tax issues, and legal issues more generally. The analysis of these items yielded two internally consistent sets of questions, namely *personal experience* (e.g., "Do you have experience in dealing with tax issues, as part of your job or studies?," 1 = *not at all*, 7 = *very much*) and *experience in family* (e.g., "Does anybody in your family have experience in dealing with tax issues, as part of their job or studies?," 1 = *not at all*, 7 = *very much*). The three items of personal experience with tax issues, business tax, and legal questions were combined into one score by averaging responses across items ($\alpha = .66$). Likewise, the three items of experience in the family with tax issues, business tax, and legal questions were combined into one score by averaging responses across items ($\alpha = .74$).

The next four pages of the questionnaire provided four examples of tax problems that the presenter solved with the participants to illustrate the meaning of the answer format for the experimental problems, and to train participants in the use of the answer sheets. This was followed by the twenty experimental problems which participants were asked to solve in the remaining time. The problems were identical for both experimental conditions. After being presented with the problem, participants were first asked to fill in eight *steps* leading to the answer to the problem. The steps naturally differed between the two sets of legislation.[7] In each condition, a further entry was required for the final *answer* ("The impact of this transaction on net income for the year ending 30 June is…"). Finally, for each problem, participants were asked whether they "followed the required steps in answering the question" (*yes/no*), how confident they were that their answer was correct (1 = *not at all*, 7 = *very much*), and how confident they were that they correctly followed the steps required by the tax law in reaching their answer (1 = *not at all*, 7 = *very much*).

Results

We first examined the comparability of the two groups for any bias arising either from prior knowledge or from the perceived difficulty of the presentation. We verified that the two experimental conditions did not differ in terms of participants' prior experience with tax and other legal issues, whether from chance factors or suboptimal randomization. *T*-tests indicated no significant

31

differences between TVM and CL conditions for measures of *personal experience*, $t(67) = 1.04$, *ns*, and *experience within the family*, $t(67) = -.089$, *ns*. *Personal experience* was generally low ($M = 1.47$), while *experience in the family* was somewhat higher ($M = 3.34$). According to their self-ratings, few participants had prior experience with tax issues (86 percent rated their experience as 1 or 2), business tax (94 percent rated experience as 1 or 2), or legal questions (88 percent rated experience as 1 or 2). These results confirmed that the randomization process had been successful and our view that students would bring with them little prior knowledge of tax issues.

The almost complete absence of experience among participants also meant that there was too little variance for this to be tested further. The research could not explore the possibly moderating effects of prior experience--there was too little experience in either group to allow us to draw any inferences about the effects of the level of prior experience.[8]

Second, we examined the results for the participants' ratings of the two presentations. The ratings of the *difficulty* and *quality* of the presentations were subjected to *t*-tests to check whether the two presentations were perceived to differ. The *difficulty* of the presentation was not rated differently in the two conditions, $t(67) = 1.50$, *ns*. In both cases participants tended to evaluate the presentation as difficult ($Ms = 4.70$ and 4.30). Ratings of the *quality* of the presentation, however, differed between the two conditions, $t(67) = -1.73$, $p =$

.089. While in either condition the *quality* of the presentations was evaluated

rather positively, the TVM presentation was considered slightly less clear, less

well-structured, or less articulate than the CL presentation (Ms = 5.05 vs. 5.46).

Even though this difference was only marginally significant, the result had

implications for our further analyses. It suggested that the quality of the

presentations was not the same between the two conditions, which could have

affected the further results. In order to control statistically for the difference,

when testing the hypotheses we included the rating of the *quality* of presentation

as a covariate.

On the other hand, it is also possible that the rating reflected the clarity

and certainty of the TVM legislation itself. It is quite conceivable that unclear

concepts simply cannot be presented as lucidly as clear concepts, and the clarity

of the TVM concept, as expressed in the legislation, is something we were

interested in. By controlling for the observed differences in the ratings of the

quality of each presentation, we might conceivably filter out part of the difference

between the legislation that the project aimed to uncover. We therefore report our

results for both analyses: without controlling for the perceived difference in the

presentations, and with the statistical control included, although in our view,

differences in the quality of the presentations are unlikely to account for our

findings.

Measures of Consistency

To test hypothesis H1 (*consistency of answer*), we investigated two indicators measuring the degree to which participants reached consensus on an answer. First, we determined for each of the twenty tax problems the answer that was shared by the greatest number of participants (i.e., the mode).[9] All answers were coded as either disagreeing (0) or agreeing (1) with the mode. For each person, we then calculated the number of tax problems where they agreed with the mode. More precisely, we averaged scores of disagreement versus agreement across tax problems, which yielded the probability of participants agreeing with the mode. We did this for each group of tax problems (that is, *problem type* A, B and C).

The agreement scores were subjected to an analysis of variance with the factors *legislation* and *problem type* (the latter as a within-subjects factor). The analysis yielded a marginally significant effect of legislation $F(1, 67) = 3.67$, $p = .060$. Contrary to the hypothesis, there tended to be less agreement on an answer under TVM compared to the CL condition ($Ms = .52$ vs. .60). That is, across all three problem types, on average, 52 percent of participants in the TVM condition agreed on the same answer and 60 percent of participants in the CL condition agreed on the same answer (see Table 1).

[INSERT TABLE 1 HERE]

The *problem type* also had an effect $F(2, 134) = 46.22$, $p < .001$, and significantly moderated the results just observed, $F(2, 134) = 7.97$, $p = .001$.

Simple effects were tested for each problem type to investigate the meaning of this interaction. The only problem type for which the *legislation* produced a significant effect was group A, $F(1, 67) = 18.99$, $p < .000$. The marginal overall effect of *legislation* was significantly due to the impact of the answers to Type A problems (*Ms* = .62 vs. .83). On average, 62 percent of participants agreed on the same answer in the TVM condition, but 83 percent agreed in the CL condition (see Table 1). So when the overall level of consensus was broken down and examined more carefully, most of the difference between the CL and TVM legislation was accounted for by the difficulties that TVM presented in answering the Type A problems. (This group included questions about how to recognize receipts and payments and how to manage the transformation from receipt to asset.)

The perceived *quality* of each presentation, when included as a covariate in the analysis, had no main effect, $F(1, 66) = .03$, *ns*, nor did it interact significantly with problem type, $F(2, 132) = 2.14$, *ns*. As a consequence the main effect for legislation and its interaction with problem type remained essentially the same.

We also used deviation from the mode as a second indicator of agreement--how much participants' responses deviated from the majority response. Because the size of the deviation would vary between tax problems as a function of the amount involved, we first *z*-standardized the responses. This guaranteed that each

task was given similar weight when levels of deviation were aggregated. The absolute difference between each (standardized) response and the respective (standardized) mode were then calculated and averaged across tax problems.[10] Again, this was done for each group of tax problem to obtain scores representing each level of *problem type*.

An analysis of variance yielded a significant difference between the two experimental conditions, $F(1, 67) = 5.66$, $p = .020$. Inconsistent with Hypothesis 1, the level of deviation from the majority answer was greater in the TVM than in the CL condition ($Ms = .68$ vs. .50). While, again, *problem type* had a significant main effect, $F(2, 134) = 13.91$, $p < .001$, the more relevant interaction effect was not significant, $F(2, 134) = 1.55$, *ns*. The latter result is inconsistent with the significant interaction effect of legislation and problem type found for the previous consensus measure. Further inspection of the data, however, showed that the only significant simple effect of legislation was again observed for tax problems of Type A, $F(1, 67) = 6.57$, $p = .013$. So, although differences between *problem types* were not as obviously pronounced for the deviation as for the consensus measure, results for the two measures converged on a similar pattern (see Table 2).

[INSERT TABLE 2 HERE]

When the perceived *quality* of the presentation was added as a covariate, this did not have a main effect, $F(1, 66) = 2.57$, *ns,* nor a significant interaction

effect with problem type, $F(2, 132) = 2.28$, *ns*. Hence, the inclusion of the covariate did not substantially affect the results.

In summary, the results for both empirical indicators of *consensus* (the same answer, or a smaller deviation among answers) contradicted Hypothesis 1-- there was generally greater consensus when participants were instructed in the current legislation than in TVM, although the greater level of *objective certainty* did seem to depend on the specific subject area. The results were somewhat ambiguous as the *consensus measure* showed a clear moderation effect of problem type, whereas the *deviation measure* did not. However, the pattern of means and simple effects for both measures were actually quite consistent: TVM lead to less objective certainty for Type A problems, though the effect was less pronounced for the other two kinds of questions. The results were not affected by perceptions about the presentation.

Measures of Confidence

Participants' confidence in the correctness of their final answer was measured on a 7-point scale (1 = *not at all*, 7 = *very much*). The ratings were averaged across the tax problems for each of the three types, ignoring missing values. An analysis of variance yielded a marginally significant effect of the experimental factor, $F(1, 67) = 3.70$, $p = .059$. Contrary to Hypothesis 2, participants in the TVM condition indicated that they were less confident about their answer than in the CL condition (Ms = 3.55 vs. 4.16). Again, *problem type*

37

had a main effect, $F(2, 134) = 35.09$, $p < .001$, but, more importantly, did not significantly interact with *legislation*, $F(2, 134) = 1.28$, *ns*. The trend of greater confidence in the correctness of one's answer under CL than TVM legislation thus held independent of *problem type* (see Table 3).

[INSERT TABLE 3 HERE]

However, the perceived *quality* of presentation included as a covariate in the analysis had a significant main effect, $F(1, 66) = 6.49$, $p = .013$; its interaction with *problem type* was not significant, $F(2, 132) = 1.14$, *ns*. The better the perceived quality of the presentation, the more confident were participants in the correctness of their answers (across *problem type*, $\beta = .28$). When the effect of the covariate was controlled, the previously significant main effect of *legislation* was no longer statistically significant, $F(1, 66) = 2.05$, *ns*. This means the trend of greater confidence in one's answers under the CL than TVM legislation was significantly related to, and thus could be due to, the perceived *quality* of the presentations. However, as discussed before, it could also be that the perceived quality of the presentation is a *consequence* of the clarity of the legislation and the confidence it elicited, and therefore should not be statistically controlled.

Participants' confidence in following the correct steps was measured by two items. First, participants indicated by *yes* or *no* whether they followed the required steps in answering the tax question. These responses were averaged across each group of tax problems, ignoring missing values when participants did

not attempt the problem. This score thus reflects the likelihood that respondents thought they followed the required steps when they attempted to solve the tax problems.

An analysis of variance yielded a statistically significant effect, $F(1, 67) = 7.34$, $p = .009$. Inconsistent with Hypothesis 3, participants reported having followed the required steps less often in the TVM than in the CL condition ($Ms = .69$ vs. $.87$). That is, participants in the TVM condition indicated in 69 percent of the cases attempted, they followed the required steps, but participants in the CL condition said so in 87 percent of the cases. Problem type had a main effect, $F(2, 134) = 14.65$, $p < .001$, but the more relevant interaction effect was only marginally significant, $F(2, 134) = 2.47$, $p < .089$. In fact, simple effects of legislation were significant (or close to significant) for all three types of tax problems, even though the effect was somewhat weaker for type A, $F(1, 67) = 3.82$, $p = .055$, and Type C problems, $F(1, 67) = 4.46$, $p = .038$, than for Type B problems, $F(1, 67) = 10.22$, $p = .002$ (see Table 4).

[INSERT TABLE 4 HERE]

Perceptions of the quality of the presentation, included as a covariate, did not have a significant main effect, $F(1, 66) = 1.90$, ns, nor a significant interaction effect with problem type, $F(2, 132) = .58$, ns. The results for legislation therefore remained unchanged when the perceived quality of each presentation was controlled.

A second item measured more explicitly participants' confidence in having followed the steps correctly on a 7-point rating scale. The ratings were again averaged across tax problems of each of the three types, ignoring missing values. An analysis of variance yielded a significant effect of legislation, $F(1, 67) = 3.97$, $p = .050$. Again, inconsistent with Hypothesis 3, participants in the TVM condition were less confident that they followed the steps correctly than in the CL condition ($Ms = 3.38$ vs. 4.01). While problem type had a significant main effect, $F(2, 134) = 33.46$, $p < .001$, the more relevant interaction effect was far from significant, $F(2, 134) = .67$, ns. Hence, the CL led to greater confidence in correctly following the required steps than TVM, irrespective of the type of tax problem (see Table 5).

[INSERT TABLE 5 HERE]

However, perceived quality of presentation treated as a covariate had a significant main effect, $F(1, 66) = 7.18$, $p = .009$, while its interaction with problem type was not significant, $F(2, 132) = .78$, ns. The better the perceived quality of the presentation, the more confident were participants that they followed the steps correctly (across problem type, $\beta = .30$). Controlling for the effect of the covariate, the previously significant main effect of legislation was no longer statistically significant, $F(1, 66) = 2.21$, ns. This means, the greater confidence in correctly following the steps under the CL than TVM *could* be due to the perceived quality of the presentations.

In summary, with regard to *subjective certainty*, contrary to Hypotheses 2 and 3, participants had greater confidence both in the correctness of their answers, and in their opinion that they had correctly followed the required steps, when instructed in the current legislation than in TVM; this was the case irrespective of problem type. While the quality of the presentation, which was rated marginally lower for TVM than current legislation, might explain the legislation effects on confidence, we consider this unlikely (and not a parsimonious explanation) given that presentation quality did not account for the findings on the other aspects of certainty reported in this article.

Measures of Accuracy

To test the subsidiary hypotheses about the accuracy of responses, we constructed a measure that reflected the probability that participants agreed with our assessment of the "correct" answer intended by the legislature. For each of the twenty tax problems, we identified one answer that we considered most likely to be the intended result. Each answer was coded as correct (1) when corresponding to the correct answer; otherwise it was coded as incorrect (0), including the case when no answer was provided. For each participant, scores of correctness were averaged across tax problems of each type. This average score reflects the probability that a participant achieved a correct response.[11]

An analysis of variance yielded a significant effect of *legislation*, $F(1, 67)$ = 7.05, $p = .010$. Inconsistent with Hypothesis 4, answers in the TVM condition

were less likely to be correct than in the CL condition (Ms = .44 vs. .53). That is, across problem types, the average probability of achieving a correct response was 44 percent under the TVM legislation but 53 percent under CL. However, this effect was clearly moderated by *problem type*; the interaction effect was statistically significant, $F(2, 134) = 27.02$, $p < .001$, while the less interesting main effect of *problem type* was also significant, $F(2, 134) = 83.81$, $p < .001$. To explore the meaning of the interaction effect, the simple effects of *legislation* for the different problem types were analyzed. There was clearly a greater probability of a correct answer under CL (83 percent) than TVM legislation (51 percent) for tax problems of Type A, $F(1, 67) = 37.34$, $p < .001$. In neither the Type B problems, $F(1, 67) = .04$, *ns*, nor the Type C problems, $F(1, 67) = .04$, *ns*, did the two experimental groups differ in the probability of finding the correct answer (see Table 6).

[INSERT TABLE 6]

The perceived *quality* of presentation, when included as a covariate, did not have any significant effects; neither a main effect, $F(1, 66) = .58$, *ns*, nor an interaction effect with *problem type*, $F(2, 132) = .03$, *ns*. Correspondingly, the results for the probability of providing the correct answer remained unchanged when the covariate was controlled.[12]

While agreement with the intended answer could indicate that participants understood and were able to apply the legislation, it could also result from

participants' intuition of the correctness or appropriateness of an answer without having fully understood or correctly applied the specific legislation. Therefore, we attempted to measure whether participants followed correctly what we considered to be the steps intended by the legislature to obtain answers to the tax problems. We coded the eight steps for each tax problem as to whether or not they corresponded in their entirety to the intended way. We then averaged scores of correctness of steps across the tax problems of each type, excluding problems that participants did not attempt to answer. This score reflected the likelihood of following the correct steps, when an answer was attempted.[13]

For most tax problems, some of the steps were irrelevant, which participants indicated by leaving the entry empty or inserting a zero. Hence, all zeros and missing entries were treated as indicating either irrelevance of the step or a nil amount. However, some participants did not attempt to answer all the tax problems, so their nonentries were not counted as correct nil amounts. Therefore, tax problems that were not attempted (that is, when none of the steps nor the final answer contained an entry), were coded as missing values and ignored. The number of tax problems participants did not attempt to answer did not differ between the two experimental conditions, $F(1, 67) = .64$, *ns*. On average, participants did not attempt to answer 2.24 tax problems in the TVM condition and 2.94 tax problems in the CL condition.

An analysis of variance for all twenty tax problems yielded a significant effect of the *legislation* factor, $F(1, 67) = 39.98$, $p < .001$. Contrary to Hypothesis 5, the likelihood of following the right steps was lower in TVM than in the CL condition (Ms = .24 vs. .47). That is, across problem types, the average probability of correctly following the steps was 24 percent under the TVM legislation but 47 percent under CL. However, *problem type* moderated this effect, as indicated by a significant interaction effect, $F(2, 134) = 79.77$, $p < .001$; the less relevant main effect of problem type was also significant, $F(2, 134) = 67.43$, $p < .001$. Simple effects of *legislation* for the different problem types illustrate the meaning of the interaction effect. There was a clearly greater probability of following the correct steps under CL (81 percent) than TVM legislation (25 percent) for tax problems of Type A, $F(1, 67) = 104.57$, $p < .001$. There was also a clearly greater probability of correct steps under CL (41 percent) than TVM (17 percent) for Type B problems, $F(1, 67) = 37.92$, $p < .001$. Conversely, a less pronounced effect for Type C problems, $F(1, 67) = 4.33$, $p = .041$, showed a greater probability of correctly following the steps under TVM (30 percent) than CL legislation (19 percent) for tasks of Type C (see Table 7).

[INSERT TABLE 7 HERE]

The perceived *quality* of the presentation, included as a covariate, had neither a significant main effect, $F(1, 66) = 2.37$, *ns*, nor a significant interaction

effect with *problem type*, $F(2, 132) = .30$, *ns*. The results for the probability of correctly following the steps were unaffected by the inclusion of the covariate.[14]

In summary, with regard to *accuracy*, the results showed that, contrary to Hypotheses 4 and 5, participants more often addressed and answered the tax problems in the intended way when instructed in the current legislation than in TVM. When the statistical tests were repeated only for those tasks that were unambiguous, the general finding was replicated, which increases our confidence in the validity of the findings. The greater *accuracy* for the current legislation compared to TVM was clearly moderated by subject area. There was greater substantive accuracy (correct answers) and procedural accuracy (correct steps) for Type A problems under the current legislation. For Type B tax problems, procedural accuracy was also greater under current legislation and substantive accuracy was significantly greater for the subset of sixteen unambiguous tasks. For Type C tax problems, there was no difference in terms of substantive accuracy but procedural accuracy was greater under TVM than current legislation.

Overall, the current legislation led to more confidence independent of problem type; it led to greater objective certainty and greater substantive accuracy for Type A problems; and it led to greater procedural accuracy for Type A and B problems. The only advantage of TVM was apparent for procedural accuracy on Type C tax problems.

Some Limitations of the Research

45

We have already mentioned one limitation of this study; namely, the restricted sample of University students. Further work would be required to check whether the findings can be generalized to other groups. It might be the case that people with more experience in tax matters (or in financial accounting, legal analysis, or business) would find TVM easier to grasp and would be better able to acknowledge its presumed advantages. Equally, one might suspect that more experienced people would find it difficult to switch to a new system, feeling confused by its unfamiliarity and inconsistencies with the known system. At least, their additional requirement to "unlearn" what they know would mean some cost during transition to a new system. The present study does not allow any conclusions on either issue. Our finding that naïve participants showed overall less certainty does suggest that TVM appears less intuitive and more complex.

A second limitation of this study, also mentioned above, arises from the choice of problems and issues in our sample; namely, the use of rather simple tax problems presented in a simplified scenario. It is obvious that real tax problems can be much richer and require taxpayers or tax professionals to formulate the issues for themselves. But, given our audience, our tax problems had to employ issues that our participants could recognize and manage. It is unclear whether the greater certainty observed for the current legislation would also hold for the more complex problems that real life generates.

A third limitation concerns the rather brief instructions in tax legislation used in this study. This not only implied that we had to be selective in the subject areas chosen for this empirical test; it also leaves the possibility that, with longer training and practice, TVM could be comprehended as well as the current legislation and, once comprehended, could lead to greater levels of certainty than the current legislation. Without further evidence, however, this too remains mere speculation.

Finally, this research tried to focus for the most part on tax problems that could be solved by applying the rules currently spelled out in the two sets of legislation. The study quite consciously did not try to explore tax issues for which there are no explicit rules in either the current legislation or the TVM legislation--the kind of problems where skill and professional judgment are called for and which often lead to litigation or require administrative clarification. It had been claimed that TVM has the advantages of providing a guiding framework from which answers to unanticipated problems might be intuitively derived, and thus one should expect that there would be fewer cases, and the need for administrative rulings and clarification would decline. Our research sheds no light on this claim.

OBSERVATIONS AND SOME IMPLICATIONS

In this section of the paper we offer some observations and tentative suggestions about the broader significance of this project.

We ask first why TVM proved not to be a superior paradigm. Several ideas seem plausible to us. First, TVM appears to require people to perform more cognitive steps in analyzing a tax problem than current law, and we take it as axiomatic that the more steps, the greater the risk of error. Looking again at mere receipts--wages and gifts--the initial analysis under either system seems equally straightforward. Under current law, wages are ordinary income; most gifts are not. Under TVM, wages are a nonprivate receipt; gifts are a private receipt. In either case, the classification can be done in one step. But this is not the end of the TVM process: wages or gifts will then sit in a drawer, be deposited into a bank account, or spent. A further step must be negotiated because TVM requires the examination of both receipts and assets; current law does not, at least not in this instance.

Second, it seems to us that TVM requires people to perform more complex cognitive reasoning processes even for very simple transactions. Other examples used above--contracts for services to be performed in the next tax year, or the sale of goods on credit--display a level of complexity that the current rules do not. Under current law, the issue is the time at which to recognize the income: the time of contract, demand for payment, or receipt. There is a simple choice between three alternatives. But TVM requires the recognition of the contract as an asset, and recognition of the obligations arising under it--it recognizes the creation of the contract rights as involving a tax event, and then requires the classification of

the contract to ascribe a tax value to the contract and obligations. This is no longer a single choice between three alternatives; there are separate choices required for the cash flows and the assets, and although the choices are independent, their effects are closely related and have to be consistent. These steps are not self-evident--nor is it realistic to think they could be accurately "guessed"; they require specialist knowledge and sophistication. Again, it is axiomatic that the more complex the cognitive tasks, the greater the risk of error.

Next, it seems to us that TVM involves a variety of counter-intuitive elements and processes. We doubt, for example, that a person depositing an amount into a bank account sees it is as the purchase of an asset. The idea that a person buys the bank account into which they deposit their cash and checks seems strained. Similarly, we doubt that a person would view the deposit of a gift as another transaction that requires separate and often divergent characterization. The transformation of cash to asset is not self-evident, and unless it is handled correctly, error is likely to follow.

More fundamentally, it seems much more likely to us that in the popular imagination individuals view income tax as being largely about cash flows and not so obviously about holdings of assets or liabilities--income is first thought of as the amounts a person receives, as wages, dividends, interest, or the sales of assets, for example (Gammie 2001). The *selling* of assets for a profit may play a part in the idea of income in the common imagination, but the annual recognition

and classification of a pool of assets is not the core of the idea. TVM makes the

recognition, classification, and valuation of assets and liabilities central to

determining taxable income and throughout the period while the asset is held,

rather than just once when the asset is sold or liability repaid. This focus on

assets and liabilities may lack the simple and deep cultural resonance of income

as a cash flow or realized gain. If this surmise is correct, it is not surprising that

TVM did not induce either greater confidence or greater accuracy--it abstracted

too far from the common understanding of what an income tax is directed toward.

This last point leads to the question of wider implications. In this article

we reported on one experiment involving a tax reform proposal, but in our view

this research has wider ramifications for public and private law projects. All

public law reform and private commercial arrangements involve choices about

alternative paradigms. At the beginning of this article we explained how the legal

rights and obligations embodied in private law regimes--mortgages and insurance

policies--and in public law regimes--environmental regulations or corporate laws-

-could be achieved using other legal paradigms. While such choices may not

always be immediately appreciated, clever reformers and practitioners will

recognize the existence of a range of options and may wish to shift to one of the

alternatives.

Our research suggests some thoughts about whether and when they would

be well or poorly advised to do so: why a new paradigm might or might not prove

an advance on current practice. Reform projects will be less than successful where a new paradigm requires people to process more or more complex cognitive steps in analyzing their legal relationships than current law. The mortgage example shows the effect of secured loan can be constructed in several ways, but it is not clear that explicitly constructing it from a call option would be more easily comprehended either by the parties or outsiders. The idea that the financier is the owner, or that amounts it receives are not interest but periodic option fees, may well be apt to mislead the occupant when other issues arise such as the expectation to insure, the liability to pay property taxes, or rights to inheritance.

In the same way, legal relationships built on counterintuitive elements and processes present a heightened risk of error. The insurance example demonstrates the point. Constructing an insurance policy from a put option is feasible, but it would suggest that the insured should have to deliver the insured property as its part of the bargain. If the property has been stolen or destroyed, that is precisely what the insured cannot do. Hence the option agreement would have to be drawn so that the insured can require the insurer to buy something the insured cannot deliver, which seems more than a little odd.

A third observation from our work is that paradigms that require substantial qualifications and adjustments will also likely face problems. Again, the mortgage example is instructive. If the mortgage were instead constructed

around an option, certain aspects of the relationship would need to be modified. For example, there would need to be a collateral agreement allowing the option-holder to occupy the property, the option agreement would need to specify that it is the occupant who has to maintain the property, and some regime would have to be created that allows the occupant to sell the property, even though it is not the owner. These modifications and elaborations may mar the attractiveness of the new paradigm to such an extent that it is not superior to the current design.

Finally, although we cannot claim to have for tested this, it seems plausible to us that some paradigms are instinctively recognized by the popular imagination, whether by habit and usage or by something inherent in the paradigm, we cannot tell. Our naïve audience does not tell us *why* they found an income tax built around flows of money easier to grasp and apply than one built around assets and liabilities, although we were able to rule out some possibilities such as a body of existing knowledge or confusion in our presentation. But one possible explanation for the subjects' preference for the current law is that there is some deep cultural resonance in the concepts used in current law. The existing structure was, consciously or unconsciously, built using the kinds of paradigms that users expected to see. When a legal relationship is built using unexpected paradigms, the dissonance causes distress and error.

Another variant of this same idea is that perhaps the repeated processes of negotiation and litigation--the invisible hand of a legal system--have led over time

to the unwitting adoption of recognized and accepted paradigms that appeal to the Western mind. We may have managed to reach a collection of preferred paradigms because alternatives that have been adopted and tested have been found wanting. Departures are, therefore, quite likely to be to suboptimal positions.

The implication of these thoughts is to suggest that there may be good reasons for reformers to prefer consciously to adopt paradigms that are already familiar, where possible. In our tax context, we doubt that personal familiarity explains our outcomes--our audience lacked experience in tax matters, and tax terms and concepts are not things that a naïve audience tends to encounter by chance. For mortgages or insurance, on the other hand, there may well be a sense of familiarity with the existing paradigm that few in the community will lack. Using other paradigms would therefore have to overcome two hurdles: unlearning what is already known, and appreciating the implications of the new paradigm.

Our research also suggests how reformers could test their proposals prior to launch. Reforms that seem obvious and desirable to their proponents do not always succeed, despite the proponents' conviction about the theory behind the measure. The public policy literature stresses the value of empirically testing complex reform packages (Dunn 1998), and may also hold some clues for the design of reforms and choice of paradigms. So far as we can tell, our experiment is the first to set out a methodology for testing the construction rather than the

language of legislation. One would hope that where a novel paradigm is proposed

for a law reform project or a private arrangement, the choice among the

competing paradigms should involve empirical assessment as well as the more

usual assertion of benefits, yet we have found no evidence that governments or

lawyers take seriously the idea that benefits claimed for particular paradigms can

and should be tested. But our research suggests that the apparent virtues of a

proposed reform can be tested and meaningful results recorded.

Until we have an established methodology, perhaps along the lines of the

experiments we undertook, and a willingness to investigate the possibilities and

trade-offs that alternative paradigms present, we will lack an important tool to

assess whether measures proposed by reformers will repay the confidence of their

proponents.

ENDNOTES

1. The board's staff produced information papers, held meetings and briefings,

commissioned an assessment of compliance costs and two projects testing TVM

(Australia Government 2002a).

[2]. The idea of "tax value" was thus not intended to be the same as market value;

that is, TVM was not meant to extend to unrealized gains or losses.

[3]. The requirement to file a Business Activity Statement is imposed on persons who earn business profits or investment income. An alternative selection criterion, disqualifying participants who had filed an annual income tax return, was considered too strict as most students would derive income from part-time jobs, requiring them to file an annual income tax return.

[4]. The commissioned external reviewers were Professor Cameron Rider, Faculty of Law, University of Melbourne and Associate Professor John Glover, Faculty of Law, Monash University.

[5]. The text of the materials prepared for the experiments can be viewed at http://www.parsons.law.usyd.edu.au/TaxationLaw/cooper.htm (last accessed February 8, 2008).

[6]. One participant was twenty-eight-years-old and thus older than the intended age limit of twenty-five. However, inclusion or exclusion of this case did not affect results. The results presented here refer to the complete sample.

[7]. In the TVM condition, the steps corresponded to the steps in the TVM formula: receipts x_1 – payments x_2 + (closing tax value of assets x_3 – opening tax value of assets x_4 = net assets x_5) – (closing tax value of liabilities x_6 – opening tax value of liabilities x_7 = net liabilities x_8). For the current law, the steps corresponded to the elements described above: ordinary income x_1 + statutory income x_2 + (capital

gains proceeds x_3 – capital gains cost base x_4 = net capital gains x_5) – deductions x_6 – specific deductions x_7 (excluding capital payments x_8).

[8]. Despite the limited range in self-reported experience, we explored the issue empirically by distinguishing between participants who stated they had no personal experience with tax issues at all (65 percent) from those participants who indicated at least a "2" on the 7-point scale of personal experience and thus some (even if minimal) experience. Analyses showed that this factor did not significantly affect or moderate the results; it was therefore dropped from the analyses presented here. Similarly, we used a median split to classify participants according to their age: 39 percent were eighteen- or nineteen-years-old, and 61 percent were between twenty- and twenty-eight-years-old. Inclusion of age category in the analyses testing our hypotheses did not yield any significant main of interaction effects of this factor; it was therefore not considered in the analyses presented here.

[9]. For one task, there were two modes in the TVM condition. One of them equaled the mode in the CL condition and was therefore used for the analyses reported here.

[10]. We ignored cases where no valid answer was provided by a participant. The number of tax problems for which participants did not offer an answer did not

differ between the two experimental conditions, $F(1, 67) = .30$, *ns*. On average, participants did not provide a final answer for 2.89 tax problems in the TVM condition and for 3.41 tax problems in the CL condition.

[11]. Four of the twenty tasks allowed for a defensible alternative to our preferred view. Ambivalent questions were deliberately included in this study in order to test the claim that TVM would result in greater consensus; for example, to see if answers to difficult questions would be evenly split under current law, but skewed to one answer under TVM. But at this stage of the analysis, we were concerned more with accuracy rather than consistency. We thus supplemented our analyses of all twenty tax problems with analyses that focused only on the sixteen tasks for which we considered the intended answer more than reasonably clear.

[12]. Analyses for the sixteen unambiguous tasks showed very similar results, namely main effects of legislation and problem type as well as a significant interaction effect. The simple effect of legislation for problem Type A was the same effect as before, as no ambiguous tasks were excluded from group A. However, there was also a significant difference for Type B problems, $F(1, 67) = 4.24$, $p = .043$. The probability of a correct answer was again greater under CL (48 percent) than under TVM legislation (38 percent). For Type C problems, there was again no significant difference between the experimental groups, $F(1,$

67) = 1.64, ns. These results were not substantially altered when perceived quality of presentation was included as a covariate.

[13]. As discussed in the previous section, for four tasks there was some ambiguity as to the correct answer and therefore also the correct steps. Again, we analyzed the data for all twenty tax problems first, and then repeated the analyses for the sixteen unambiguous tasks.

[14]. The analyses for the subset of sixteen unambiguous tasks yielded very similar results.

REFERENCES

Alm, James, and Sarah Jacobsen. 2007. Using Laboratory Experiments in Public

   Economics. *National Tax Journal* 60:129-52.

Australian Government. 1999. New Business Tax System – Stage 2 Response.

   Press Release No 74/1999, November 11.

   http://www.treasurer.gov.au/tsr/content/pressreleases/1999/074.asp (last

   accessed August 20, 2008).

Australian Government. 2000. Tax Value Method. *Press Release No 81/2000*, 11

   November, 1999. Available at

   http://www.treasurer.gov.au/tsr/content/pressreleases/2000/081.asp (last

   accessed August 20, 2008).

Australian Government. 2002a. Tax Value Method Final Recommendation.

   http://www.taxboard.gov.au/content/tvm_index.asp (accessed July 29, 2008)

Australian Government. 2002b. TVM Consultative Briefing – 6 March 2002.

   http://www.taxboard.gov.au/content/Tax_Value_Method/tvm_consultative_br

   iefing/index.asp (accessed July 29, 2008).

Brealey, Richard A., Stewart C. Myers, and Franklin Allen. 2006. *Principles of*

   Corporate Finance, 8th ed. New York: McGraw-Hill/Irwin.

Charrow, Robert P., and Veda R. Charrow. 1979. Making Legal Language

   Understandable: A Psycholinguistic Study of Jury Instructions. *Columbia Law*

   Review 79:1306-74.

Cronan, John P. 2002. Is Any of This Making Any Sense: Reflecting on Guilty

  Pleas to Aid Criminal Juror Comprehension. *The American Criminal Law

  Review* 39:1187-1259.

Diamond, Shari Seidman, and Judith N. Levi. 1996. Improving Decisions on

  Death by Revising and Testing Jury Instructions. *Judicature* 79:224-32.

Dunn, William N. 1998. *The Experimenting Society: Essays in Honor of Donald

  T. Campbell*. New Brunswick, NJ: Transaction Press.

Elwork, Amiram, Bruce D. Sales, and James J. Alfini. 1982. *Making Jury

  Instructions Understandable*. Charlottesville, VA: Michie.

Fleischer, Victor. 2002. Deals: Bringing Corporate Transactions into the Law

  School Classroom. *Columbia Business Law Review* 2002:475-97.

Gammie, Malcolm. 2001. *TVM, the Tax Base and its Relationship to Commercial

  Accounting Methods.* Paper presented at the Tax Value Method Consultative

  Conference, July 23-24, in Sydney, Australia.

Garvey, Stephen P., Sheri Lynn Johnson, and Paul Marcus. 2000. Correcting

  Deadly Confusion: Responding to Jury Inquiries in Capital Cases. *Cornell

  Law Review* 85:627-56.

Genn, Hazel, Martin Partington, and Sally Wheeler. 2006. *Law in the Real World:

  Improving our Understanding of How Law Works*. London: Nuffield

  Foundation.

Gilson, Ronald J. 1987. Value Creation by Business Lawyers: Legal Skills and Asset Pricing. *Yale Law Journal* 94:239-313.

Halperin, Daniel. 1974. Business Deductions for Personal Living Expenses: A Uniform Approach to an Unsolved Problem. *University of Pennsylvania Law Review* 122:859-933.

Klein, William A. 1966. The Deductibility of Transportation Expenses of a Combination Business and Pleasure Trip--A Conceptual Analysis. *Stanford Law Review* 18:1099-1118.

Kling, Jeffrey R. 2007. Methodological Frontiers of Public Finance Experiments. *National Tax Journal* 60:109-27.

Margolis, Kenneth R. 1998. Responding to the Value Imperative: Learning to Create Value in the Attorney-Client Relationship. *Clinical Law Review* 5:117-77.

May, Christopher N. 1995. "What Do We Do Now?": Helping Juries Apply the Instructions. *Loyola of Los Angeles Law Review* 28:869-902.

Mellinkoff, David. 1983. The Myth of Precision and the Law Dictionary. *UCLA Law Review* 31:423-42.

Posner, Richard A. 2003. *Economic Analysis of Law*, 6th ed. New York: Aspen Publishers.

Power, Robert C. 1999. Reasonable and Other Doubts: The Problem of Jury Instructions. *Tennessee Law Review* 67:45-124.

Review of Business Taxation. 1999a. *A Tax System Redesigned: More Certain, Equitable and Durable*. Canberra: Treasury.

Review of Business Taxation. 1999b. *A Tax System Redesigned: Draft Legislation*. Canberra: Treasury.

Review of Business Taxation. 1999c. *A Tax System Redesigned: Explanatory Notes*. Canberra: Treasury.

Samansky, Allan J. 1981. Hobby Loss or Deductible Loss: An Intractable Problem. *University of Florida Law Review* 34:46-71.

Smith, Clifford W., Jr. 1990. Applications of Option Pricing Analysis. In *The Modern Theory of Corporate Finance*, 2nd ed., ed. Clifford W. Smith. New York: McGraw-Hill/Irwin.

Steele, Walter W. Jr., and Elizabeth G. Thornburg. 1988. Jury Instructions: A Persistent Failure to Communicate. *North Carolina Law Review* 67:77-120.

University of Illinois Law Review. 2002. Symposium: Empirical and Experimental Methods in Law. University of Illinois Law Review 2002:789-1176

STATUTES CITED:

Income Tax Assessment Act 1997.

http://www.austlii.edu.au/au/legis/cth/consol_act/itaa1997240/ (accessed August, 20, 2008).

TABLE 1.

Objective Certainty – Consensus Measure (Estimated Means)

	Legislation	
	TVM	CL
Problem Type		
Problem Type A	.62	.83***
Problem Type B	.49	.52
Problem Type C	.46	.44
Overall	.52	.60

Note: According to simple effect analyses, means for TVM and CL conditions differ at

*** $p < .001$.

ANOVA results: Legislation, $F(1, 67) = 3.67$, $p = .060$; Problem Type, $F(2, 134) = 46.22$, $p < .001$; Legislation X Problem Type, $F(2, 134) = 7.97$, $p = .001$.

TABLE 2.

Objective Certainty-- Deviation Measure (Estimated Means)

	Legislation	
	TVM	CL
Problem Type		
Problem Type A	.53	.22
Problem Type B	.71	.65
Problem Type C	.80	.62
Overall	.68	.50

Note: ANOVA results: Legislation, $F(1, 67) = 5.66$, $p = .020$; Problem Type, $F(2, 134) = 13.91$, $p < .001$; Legislation X Problem Type, $F(2, 134) = 1.55$, *ns*.

TABLE 3.

Subjective Certainty – Confidence in Correctness of Answer (Estimated Means)

	Legislation	
	TVM	CL
Problem Type		
Problem Type A	4.16	4.81
Problem Type B	3.29	4.11
Problem Type C	3.19	3.58
Overall	3.55	4.16

Note: ANOVA results: Legislation, $F(1, 67) = 3.70$, $p = .059$; Problem Type, $F(2, 134) = 35.09$, $p < .001$; Legislation X Problem Type, $F(2, 134) = 1.28$, *ns*.

TABLE 4.

Subjective Certainty – Having Followed the Required Steps (Estimated Means)

	Legislation	
	TVM	CL
Problem Type		
Problem Type A	.80	.91†
Problem Type B	.65	.89**
Problem Type C	.62	.79*
Overall	.69	.87

Note: According to simple effect analyses, means for TVM and CL conditions differ at

$^{\dagger} p < .1,$ $^{**} p < .01,$ $^{***} p < .001.$

ANOVA results: Legislation, $F(1, 67) = 7.34$, $p = .009$; Problem Type, $F(2, 134) = 14.65$, $p < .001$; Legislation X Problem Type, $F(2, 134) = 2.47$, $p = .089$.

TABLE 5.

Subjective Certainty--Confidence in Correctly Following the Steps

(Estimated Means)

	Legislation	
	TVM	CL
Problem Type		
Problem Type A	4.01	4.59
Problem Type B	3.12	3.93
Problem Type C	3.00	3.51
Overall	3.38	4.01

Note: ANOVA results: Legislation, $F(1, 67) = 3.97$, $p = .050$; Problem Type, $F(2, 134) = 33.46$, $p < .001$; Legislation X Problem Type, $F(2, 134) = .67$, *ns.*

TABLE 6.

Accuracy--Agreement with Intended Answer (Estimated Means)

	Legislation	
	TVM	CL
Problem Type		
Problem Type A	.51	.83***
Problem Type B	.43	.42
Problem Type C	.36	.35
Overall	.44	.53

Note: According to simple effect analyses, means for TVM and CL conditions differ at

*** $p < .001$.

ANOVA results: Legislation, $F(1, 67) = 7.05$, $p = .010$; Problem Type, $F(2, 134) = 83.81$, $p < .001$; Legislation X Problem Type, $F(2, 134) = 27.02$, $p < .001$.

TABLE 7.

Accuracy--Following the Correct Steps (Estimated Means)

	Legislation	
	TVM	CL
Problem Type		
Problem Type A	.25	.81***
Problem Type B	.17	.41***
Problem Type C	.30	.19*
Overall	.24	.47

Note: According to simple effect analyses, means for TVM and CL conditions

differ at

* $p < .05$, *** $p < .001$.

ANOVA results: Legislation, $F(1, 67) = 39.98$, $p < .001$; Problem Type, $F(2, 134)$

$= 67.43$, $p < .001$; Legislation X Problem Type, $F(2, 134) = 79.77$, $p < .001$.