



SCHOOL OF LAW
TEXAS A&M UNIVERSITY

Texas Wesleyan Law Review

Volume 11 | Issue 2

Article 11

3-1-2005

The Common Law of Contracts: Are Broad Principles Better Than Detailed Rules? An Empirical Investigation

M. P. Ellinghaus

E. W. Wright

Follow this and additional works at: <https://scholarship.law.tamu.edu/txwes-lr>

Recommended Citation

M. P. Ellinghaus & E. W. Wright, *The Common Law of Contracts: Are Broad Principles Better Than Detailed Rules? An Empirical Investigation*, 11 Tex. Wesleyan L. Rev. 399 (2005).

Available at: <https://doi.org/10.37419/TWLR.V11.I2.10>

This Symposium is brought to you for free and open access by Texas A&M Law Scholarship. It has been accepted for inclusion in Texas Wesleyan Law Review by an authorized editor of Texas A&M Law Scholarship. For more information, please contact aretteen@law.tamu.edu.

THE COMMON LAW OF CONTRACTS: ARE BROAD PRINCIPLES BETTER THAN DETAILED RULES? AN EMPIRICAL INVESTIGATION¹

M.P. Ellinghaus[†] and E.W. Wright[‡]

I.	INTRODUCTION	399
II.	THE THREE LAW MODELS	401
	A. <i>Case Law</i>	402
	B. <i>UPICC</i>	402
	C. <i>ACC</i>	403
III.	ILLUSTRATION: “BASE METALS V. PRECIOUS METALS”	403
	A. <i>Case Law Relevant to Base Metals v. Precious Metals</i>	404
	B. <i>UPICC Provisions Relevant to Base Metals v. Precious Metals</i>	405
	C. <i>ACC Provisions Relevant to Base Metals v. Precious Metals</i>	407
IV.	METHODOLOGY	408
	A. <i>Experimental Paradigm</i>	408
	B. <i>Design</i>	409
	C. <i>Materials</i>	410
	D. <i>Data</i>	410
V.	RESULTS	411
	A. <i>Predictability</i>	411
	B. <i>Justice</i>	413
	C. <i>Accessibility</i>	418
	D. <i>Efficiency</i>	419
VI.	CONCLUSION	420

I. INTRODUCTION

The common law of contracts consists of precepts of varying levels of determinacy stated in non-canonical form² in judgments given in the course of deciding real life disputes. It includes, at one end of a

1. This Article is substantially based on research conducted with the support of the Australian Research Council (Linkage Program) and the Law and Justice Foundation of New South Wales. A full report is soon to be published as M.P. Ellinghaus and E.W. Wright (with M. Karras), *Models of Contract Law: An Empirical Evaluation of Their Utility*, L. & JUST. FOUND. OF N.S.W. (forthcoming).

[†] Law School, University of Melbourne, Australia.

[‡] School of Law, University of Newcastle, Australia.

2. The term “canonical” is used by FREDERICK SCHAUER, *PLAYING BY THE RULES A PHILOSOPHICAL EXAMINATION OF RULE-BASED DECISION-MAKING IN LAW AND LIFE* 12–13, 181 (1991).

spectrum, rules that apply when relatively few facts are established. At the other end of the spectrum there are broad principles, the application of which requires wider-ranging investigation of the facts.

The appropriate level of determinacy of common law rules is the subject of a long-standing jurisprudential debate about the relative utility of "rules" and "principles." The debate includes controversy about the meaning of these two terms. We have adopted the rough-and-ready dichotomy of "broad principles" and "detailed rules" in the hope of transcending the established theoretical discourse of principles, standards, and rules, and their many derivatives and embellishments.

The greater level of indeterminacy that is a necessary feature of broad principles typically arouses concern among common lawyers, whose natural tendency is to mediate their application by elaborating ancillary and more detailed rules which, it is hoped, will deprive broad principles of open-ended operation and reduce the scope of forensic inquiry. This is particularly true in the law of contracts.

The debate about the relative utility of broad principles and detailed rules underlies a number of recurrent controversies of particular relevance to the enforcement of contract law. One example of these recurrent controversies is the ambit of broad standards of conduct such as good faith and unconscionability. The adoption of such standards is often characterised as pursuing pragmatism instead of principle, or abandoning general justice in favour of individualised dispute resolution.³

This debate is also at the core of the current controversy in Australia and the UK over discretionary remedialism, that is, the view that courts should have discretion to award any remedy appropriate in an individual case, rather than being limited to specified remedies. Opponents of remedial discretion base their argument on the view that to discard specification in favour of discretion is to descend into an intuitive and unpredictable form of justice that ultimately involves abandonment of the rule of law. Proponents of discretionary remedialism, on the other hand, regard it as "a valid and valuable part of legal decision-making."⁴

The case against the use of broad principles is that broad principles replace the finely tuned logic and specificity of detailed rules with generalized discretions, leading to unpredictable outcomes.⁵ It is said that

3. See P.S. Atiyah, *From Principles to Pragmatism: Changes in the Function of the Judicial Process and the Law*, 65 IOWA L. REV. 1249 (1980); A. Murray Gleeson, *Individualised Justice—The Holy Grail*, 69 AUSTL. L.J. 421 (1995).

4. Simon Evans, *Defending Discretionary Remedialism*, 23 SYDNEY L. REV. 463, 464 (2001). For a general account of this controversy, see Darryn M. Jensen, *The Rights and Wrongs of Discretionary Remedialism*, 2003 SING. J. LEGAL STUD. 178.

5. Arthur T. von Mehren, *Some Reflections on Codification and Case-Law in the Twenty-First Century*, 31 U.C. DAVIS L. REV. 659, 667 (1998); cf. Jack Beatson, *Has the Common Law a Future?* 56 CAMBRIDGE L.J. 291, 299–300 (1997).

detailed rules lead to more predictable and more just outcomes, with attendant gains in accessibility and efficiency. This view has dominated the development of the common law for at least two centuries. Most readers will be familiar at least in outline with this jurisprudential debate and will know their own position on it.⁶

In this Article we are not endeavouring to engage in the debate at the level of theory. Rather, we are responding to the absence of any empirical verification of the assumptions on which the theory proceeds. Whether broad principles or detailed rules produce more just or more predictable outcomes is a question that has largely been left to the realm of speculative analysis. Additionally, no empirical studies have been conducted despite commentators' occasional requests for them.⁷

In this Article we report the results of three experiments involving the participation of 1800 subjects (law students and non-law students) in the resolution of disputes and the evaluation of judgments, using three different law models: (1) the common law of contracts (Case Law); (2) *UNIDROIT Principles of International Commercial Contracts*, a model code published by the International Institute for the Unification of Private Law in 1994 (UPICC); (3) the *Australian Contract Code*, a model code, written by us, and published by the Law Reform Commission of Victoria in 1992 (ACC). The research on which this Article is based was prompted by our interest in the codification of Australian contract law. In order to demonstrate its relevance to the jurisprudential debate, it is necessary to give some description of each of these three models.

II. THE THREE LAW MODELS

All three models address the same basic issues of contract law: how a contract is made, how its content is determined, when its performance is excused, and what remedies are available for its breach. A detailed comparison of their doctrinal content is not necessary. There are some interesting minor differences,⁸ but such differences are unlikely to affect utility. However, the three models differ markedly in the form in which the rules are expressed, their total number, and their level of detail. In our view, these are the major points of difference affecting their utility.

6. Hugh Collins, *Transaction Costs and Subsidiarity in European Contract Law*, in AN ACADEMIC GREEN PAPER ON EUROPEAN CONTRACT LAW 269, 280 (Stefan Grundmann & Jules Stuyck eds., 2002).

7. E.g., Evans, *supra* note 4, at 494–95.

8. For example: unlike Australian Case Law, UPICC does not require “consideration” as an element of formation of a contract; it has no “parol evidence” rule excluding extrinsic evidence in relation to contract documents; it imposes a substantive limit on terms limiting or excluding liability for nonperformance; it confers a right to cure a breach; it imposes a duty to renegotiate in cases of hardship.

A. Case Law

As stated at the outset, the rules of Case Law are found in the published judgments of the courts and stated in non-canonical form. Their total number is large, unknown, and, on one view of precedent, endlessly growing. Case Law includes both broad principles and detailed rules. Although Case Law does include a number of broad principles, these are not applied directly in determining rights and obligations. Their application is mediated by many detailed rules.

Several recent decisions from the High Court of Australia dealing with the unconscionable exercise of contractual rights to terminate for failure to perform on time provide a striking illustration. The High Court has made it clear that, in determining whether such a right has been exercised unconscionably, the courts are not entitled simply to apply the standard of conscience to particular facts, but must apply established doctrines that define its legal operation. In a recent case, *Tanwar Enterprises Pty Ltd. v. Cauchi*,⁹ the court said:

The terms “unconscientious” and “unconscionable” . . . describe in their various applications the formation and instruction of conscience by reference to well developed principles It is to those principles that the court has first regard rather than entering into the case at that higher level of abstraction involved in notions of unconscientious conduct in some loose sense where all principles are at large The conscience . . . which equity seeks to relieve, is a “properly formed and instructed conscience.”¹⁰

B. UPICC

UPICC is a code of law applicable to international commercial contracts promulgated by the International Institute for the Unification of Private Law (UNIDROIT), an independent organization funded by 59 member states, including both Australia and the United States of America. It has been drafted as a model uniform law suitable for domestic as well as international contracts. Many jurisdictions have applied the UPICC in arbitration proceedings even though no jurisdiction has adopted the UPICC as its law.

When we conducted our research, UPICC had 119 Articles. Since then, there have been additions, which have brought the total to 194 Articles. The UPICC begins with a chapter of General Provisions containing broad principles. These provisions include a principle imposing a duty to act in accordance with good faith and fair dealing which the parties may not exclude. The following nine chapters of rules deal in detail with formation, validity, interpretation, content, performance and non-performance, third party rights, set-off, assign-

9. (2003) H.C.A. 57.

10. *Id.* at ¶¶ 20–22 (citing *Austl. Broad. Corp. v. Lenah Game Meats Pty Ltd.*, (2001) 208 C.L.R. 199, 227 (Austl.) (Gleeson, C.J.)).

ment, and limitation periods. The UPICC provides numerous black letter rules of varying levels of detail on all of these topics of contract law. It also elaborates these rules further by the use of lengthy comments and illustrations.

C. ACC

The Victorian Law Reform Commission published the ACC in 1992. The ACC consists of only 27 Articles. There is a commentary of 47 short paragraphs; however, the ACC contains no illustrations. Unlike Case Law and the UPICC, the ACC consists entirely of broad principles. For example, the formation of a contract is covered in three Articles, the determination of its content in four Articles, and excuses from performance in three Articles.

Article 27 of the ACC explicitly overrides all other provisions in the ACC. A person may not assert a right or deny an obligation to the extent that it would be unconscionable to do so.¹¹ The commentary explains that “unconscionable” means “offending against conscience” when “judged by reference to both the values of the wider community and . . . the particular environment,” but adds that “it is impossible to specify or define exhaustively the circumstances in which it is unconscionable to assert or deny a right by the elaboration of particular rules.”¹²

Additionally, Article 3 of the ACC is designed to prevent the accretion of mediating detailed rules.¹³ Neither past nor future decisions govern the application of the Code.¹⁴ One will see that the ACC’s reliance on broad principles could not be much more overt or complete.

III. ILLUSTRATION: “*BASE METALS V. PRECIOUS METALS*”

An example drawn from our experimental materials demonstrates the difference in the use of detailed rules and broad principles, between Case Law and UPICC on the one hand and the ACC on the other. This document is based on a reported case in an Australian appellate court, which we have renamed *Base Metals v. Precious Metals*.

The case concerns two mining companies, which executed a document described as a “heads of agreement.” This permitted Base Metals to explore a mining tenement owned by Precious Metals. The document contained only six clauses, which specified payments to be made and money to be spent on exploration activity, and dealt with

11. M.P. ELLINGHAUS & E.W. WRIGHT, AN AUSTRALIAN CONTRACT CODE art. 27 (1992) [hereinafter ACC].

12. *Id.*

13. *Id.* art. 3.

14. *Id.*

some other matters touching on the relationship between the parties. The final clause of the document stated: "The above forms a heads of agreement which constitutes an agreement in itself intended to be replaced by a fuller agreement not different in substance or form."

Base Metals made the required payments and expended the required money on exploration, and discovered a valuable deposit of base metals. The parties simultaneously negotiated towards a fuller agreement, but disagreed on a number of subsidiary points.

Precious Metals was taken over by new owners. The new owners terminated negotiations. They claimed that the heads of agreement did not constitute a binding contract because the parties had not intended to make one, and because they had not defined their obligations with sufficient certainty. Base Metals sued Precious Metals for breach of contract.

The three law models agree broadly in their response to such a claim. Each of them holds that a contract is formed only if the parties intend to be legally bound, and if their agreement is sufficiently certain. Each also provides rules for supplementing the express obligations of a contract by implication.

A. *Case Law Relevant to Base Metals v. Precious Metals*

The most relevant rules of Case Law, as applied in the original case, are set out below:

General Principle

1. There is no binding contract unless the parties intend to create legally binding relations.

Mediating Rules

1. The intention of the parties is determined objectively by reference to what a reasonable person would infer.
2. Evidence of their subjective intention is inadmissible.
3. Where important terms are uncertain, lack of intention to be legally bound can be inferred.
4. Where the parties contemplate the execution of a formal contract the case may belong to The Victorian Law Reform Commission published the ACC in 1992 one of four classes:
 - (a) They intend to be bound immediately. While they expect to sign a formal contract they do not promise to do so.
 - (b) They intend to be bound immediately and promise to sign a formal document not different in effect.
 - (c) They intend to be bound immediately but performance is conditional on signature of a formal document.
 - (d) They intend not to be bound unless they sign a formal contract.

General Principle

1. There is no binding contract if essential terms of the contract are missing or uncertain.

Mediating Rules

1. A term is not uncertain simply because it has more than one possible meaning.
2. A term is uncertain only if its language is so obscure that the court cannot attribute any particular intention.
3. There is no binding contract where an *essential* term is left to be settled by future agreement.
4. But there is no obstacle to the parties leaving *important* terms to be agreed later.
5. What is essential depends on the intention of the parties.
6. There is an implied obligation to do all such things as are necessary to enable the other party to have the benefit of the contract.
7. Where the parties have agreed to do something, which cannot be done unless both concur in doing it, it is implied that they will do what is necessary.

The court might have referred to further detailed rules governing implication of terms:

1. Where the contract belongs to an identifiable class of contracts, a term may be implied if it:
 - (a) is “necessary to prevent rights from being rendered nugatory,” or
 - (b) accords with “perceived necessities of the times,” or
 - (c) is customary.
2. A term may be implied ad hoc, if it:
 - (a) is necessary for business efficacy,
 - (b) reasonable and equitable,
 - (c) obvious,
 - (d) capable of clear expression, and
 - (e) does not contradict any express term.

As can be seen, there are in fact two general principles, but their operation is mediated by a number of detailed rules which most readers will recognise. They will also appreciate that the potential for endless elaboration of these rules is an inherent feature of the Case Law method. For example, a current matter of controversy exists in Australian law whether, where the parties contemplate the execution of a formal contract, the case may belong to one of three classes—Case Law, the UPICC, or the ACC—rather than four, as stated above.¹⁵

B. *UPICC Provisions Relevant to Base Metals v. Precious Metals*

It would be impracticable to set out verbatim all the rules in UPICC that are relevant to this dispute within the format of this Article; however, they are summarised below:

15. Elisabeth Peden et al., *When Three Just Isn't Enough: The Fourth Category of the 'Subject to Contract' Cases*, 20 J. CONT. L. 156 (2004).

1. Intention to be bound and sufficient definiteness of content are necessary elements of a contract formed by offer and acceptance: Articles 2.1, 2.2.¹⁶
2. A contract may also be concluded "by conduct . . . sufficient to show agreement": Article 2.1. (It is not clear whether intention to be bound and sufficient definiteness are also prerequisites of formation by conduct, although it would be strange if they were not.)¹⁷
3. Whether there has been conduct sufficient to show agreement must be decided in accordance with criteria of intention, meaning, and regard for circumstances, set out in three separate Articles, 4.1, 4.2, and 4.3, containing ten paragraphs or subparagraphs.¹⁸
4. If a party insists that the contract is not concluded until there is agreement on specific matters or in a specific form, then no contract is concluded until that stipulation is satisfied: Article 2.13.¹⁹
5. Where a party makes it clear that it does not intend to be bound unless a formal document is drawn up, there is no contract until this is done: Article 2.13, Comment 2.²⁰
6. As a general rule a contract is concluded if parties reach an agreement on essential terms, while minor terms may be implied: Article 2.13, Comment 1.²¹
7. UPICC provides a plurality of sources for the supplementation of express terms:
 - (a) Each party has a duty to act in accordance with good faith and fair dealing: Article 1.7.²²
 - (b) Each party has a duty not to act inconsistently with an understanding which it has caused the other party to have: Article 1.8.²³
 - (c) Each party must cooperate when this may reasonably be expected: Article 5.3.²⁴
 - (d) Article 5.2 states that implied obligations may stem from
 - (i) the nature and purpose of the contract;
 - (ii) practices and usages;
 - (iii) good faith and fair dealing; and
 - (iv) reasonableness.²⁵

16. UNIDROIT PRINCIPLES art. 2.1-2.2 (1994).

17. *Id.* art. 2.1.

18. *Id.* art. 4.1-4.3.

19. *Id.* art. 2.13.

20. *Id.* art. 2.13 cmt. 2.

21. *Id.* art. 2.13 cmt. 1.

22. *Id.* art. 1.7.

23. *Id.* art. 1.8. This Article was added to UPICC after the conclusion of our experimental work, and was not included in the law statements used in the experiments.

24. *Id.* art. 5.3.

25. *Id.* art. 5.2.

- (e) Where the parties have omitted a term which is “important for a determination of their rights and duties” an appropriate term must be supplied: Article 4.8.²⁶
- (f) What is appropriate must be determined with regard to the intention of the parties, nature and purpose of the contract, good faith and fair dealing, and reasonableness.

As can be seen, UPICC also has many detailed rules, which in some respect resemble those of Case Law, but in other respects introduce further complexities; for example, the need to distinguish between “implied” and “supplied” terms.

C. ACC Provisions Relevant to *Base Metals v. Precious Metals*

The ACC provides five concise Articles relevant to the dispute. They may be quoted verbatim:

1. Article 5: A contract is made only when the parties intend legal obligations to arise.²⁷
2. Article 7: There is no contract if a necessary term is missing, is too vague, or has been left to future agreement. (The Commentary to Article 7 states a term is necessary if, without it, it is not possible to discern what each party was intended to get from their promises, and that missing detail can be supplied under Article 10.)²⁸
3. Article 10.2: The obligations of the parties are to do everything which conscience requires to ensure that each gets the benefit intended by their promises.²⁹
4. Article 26: A person who makes an assumption of any kind may require another person to act in accordance with that assumption to the extent that it would be unconscionable not to do so.³⁰
5. Article 27: A person may not assert a right or deny an obligation to the extent that it would be unconscionable to do so.³¹

The ACC provides no mediating rules governing the application of these broad principles. It does not provide rules for ascertaining the parties’ intentions or interpreting their statements or conduct. It provides only a single source for the supplementation of express content. It does not provide specific rules for agreements contemplating the preparation of a formal document.

The contrast between Case Law and UPICC on the one hand and the ACC on the other, which so clearly emerges from their application to *Base Metals v. Precious Metals*, demonstrates that an investigation of their relative utility is at the same time an investigation of both the

26. *Id.* art. 4.8.

27. ACC art. 5.

28. *Id.* art. 7.

29. *Id.* art. 10.2.

30. *Id.* art. 26.

31. *Id.* art. 27.

relative utility of broad principles and of detailed rules. Therefore, whenever the results indicate that Case Law and UPICC are superior to the ACC, this can be interpreted as evidence of the greater utility of detailed rules. Whenever the results indicate that the ACC is superior to UPICC and Case Law, one can interpret this as evidence of the greater utility of broad principles.

IV. METHODOLOGY

A. *Experimental Paradigm*

Our research used an experimental paradigm. Although most readers will understand in a general way what this means, it is only rarely encountered in legal research. It is therefore desirable, before descending into the detail of our research design that some attention is given to some special attributes of this paradigm.

As we have pointed out, the debate about detailed rules and broad principles involves a number of assertions about their relative utility. For example, one can claim that detailed rules lead to more predictable decisions. However, it is not possible to evaluate this claim simply by observing legal decision-makers in their day-to-day activities as would, say, a naturalist interested in the adaptive qualities of certain bill shapes observe the behavior of birds. There are a myriad of factors that may affect the outcomes of particular decisions, for example, the subject matter and normative content of the rules, the factual details of a dispute, and the characteristics of the decision maker, including what she ate for breakfast. It would be impossible to isolate, by natural observation, the effect of detailed rule or broad principle from the myriad of other factors; however, this does not necessarily mean that the effect of form is not important.

An experimental paradigm is designed to address the problem of isolating the effect of one of many possible variables. The effect of the research variable of interest can be isolated by independently manipulating it and measuring the effects on dependent variables, while keeping all other potential variables as equal as possible. For example, the effect of the law model on decision making can be isolated by controlling some other variables (e.g., by using the same dispute repeatedly) and randomizing others (e.g., by assigning participants to law model groups in a repeated sequence, thereby ensuring approximately equal numbers of muesli and bacon-and-egg eaters in each).

Admittedly, experiments must generally be conducted in conditions which are only abstractions of real world conditions. But this does not make the results "unreal." The method yields at least one very real result: whenever a significant difference between experimental groups is observed, one can say, definitively, that in these conditions the experimental variable caused this difference. The further significance of

this conclusion then, of course, depends on the extent to which the experimental conditions reasonably approximate the conditions of the real world. We believe that our experiments approximated real world conditions sufficiently to make the results interesting.

B. *Design*

In Experiment 1, law students enrolled in Contract Law were given the facts of a contract dispute and a statement of the relevant law drawn from one of the three law models, and were asked to decide the dispute. Ten different disputes were used, involving the full range of contract issues. One of them (*Base Metals v. Precious Metals*) has been described above. As there were three models, and ten different disputes, there were thirty experimental groups. We assigned ten students to each. Thus, three hundred student judges participated in Experiment 1.

Experiment 2 was a replication of Experiment 1, with one difference: the student judges worked in pairs. A concern existed about how diligently the single judges would apply themselves to their task; it was, after all, like sitting for an exam, without the incentive of being graded. We hoped pairs would feel accountable to each other, resulting in greater diligence and less idiosyncratic application. Six hundred students participated in Experiment 2.

In Experiment 3, non-law students were asked to read and evaluate two judgments applying one of the law models, one for the plaintiff and one for the defendant. The same ten disputes were used. In order to distinguish the effect of the law model from the effect of outcome, every possible pair-combination of judgment was administered with equal frequency, and the order in which the pairs were presented was counterbalanced. As there are nine different possible pairs of six different judgments (three law models \times two judgments) there were ninety experimental groups. We assigned ten students to each of these. In all, nine hundred students participated in Experiment 3.

It might be thought that law students do not sufficiently resemble the qualified lawyers and judges of the real world. Of course, law students were employed for reasons of sheer practicality. However, law students share many attributes of practising lawyers, particularly in Australia, where contract law students usually have completed three other law subjects in the previous two years of their university studies. In any event, if the thought is that law students have more in common with non-lawyers, it should be remembered that an evaluation of the utility of law must take into account not only the perspective of lawyers, but also that of non-lawyers who must use the law in the real world.

C. *Materials*

In Experiments 1 and 2, participants were given a written set of instructions, a statement of facts for the dispute they had to decide, a statement of the relevant law (drawn from the law models), a form on which to record their decision with reasons, and a questionnaire. The ten disputes were selected according to a number of criteria. In order to ensure that the results were not peculiar to only a particular compartment of the law, the ten disputes covered the full spectrum of contract doctrine. Each was based on a reported Australian appellate court case. We chose only split decisions³² in order to ensure that the decision would be a difficult one. We hoped this would cause students to feel a real need to use the legal materials. For the purposes of Experiment 3, we also wanted a decision for either party to be at least plausible. Additionally, in order to ensure that the dispute was not familiar to the students, we chose only cases that were not referred to in any detail in the standard student texts and casebooks.

The statements of facts were prepared based on the reported judgments in the original case. They ranged in length from one to four pages. Statements of the relevant UPICC and ACC provisions, quoting from the text of their articles and their commentary, were also provided as part of the experiments. Statements of Case Law were prepared by drawing on the actual language of both the majority and minority judgments. By preparing these statements, we removed much of the lead from the saddlebags of the common law as this made it unnecessary for users to locate, read, and synthesize the relevant authorities. While in part this was done for practical reasons, it also means that the results are more likely to be attributable to the differences of form between the law models described above. The statements of law also ranged in length from one to four pages.

The Case Law judgments were abridgements of the original cases. The UPICC and ACC judgments applied their relevant articles and commentary; further, an attempt was made to preserve the conventions of judicial style. So far as this could be done, the reasons in all the judgments referred to the same factual grounds for decision. The judgments ranged in length from three to six pages.

D. *Data*

The dependent variables were designed to be measures of utility. In keeping with the empirical nature of the study, we adopted a conception of utility, which reflected the demands most often made of legal systems by lawyers and non-lawyers alike. These demands were that the law should be:

32. Five of the original majority opinions were in favour of the plaintiff. The other five were for the defendant.

1. Certain (produce predictable outcomes)
2. Just (produce fair outcomes)
3. Accessible (use clear language and logic)
4. Efficient (easy to locate, comprehend, and apply)

In Experiments 1 and 2, the data we collected included the student judges' decisions, their reasons (they were specifically directed to indicate how they applied the law), and the time in which they finally reached their decision. Their evaluations of the law were measured by their responses to eight propositions on a seven-point scale of agreement or disagreement (the Likert technique). The propositions related to the aspects of utility identified above. In Experiment 3, the participants were asked to complete two identical questionnaires. Each questionnaire asked the students to evaluate the judgment they had just read. Again, the Likert technique was employed using nine propositions.

V. RESULTS

A. *Predictability*

Level of agreement is a measure of predictability or certainty. The level of agreement was measured by taking the number of decisions in favour of plaintiff and defendant, and expressing the majority to minority ratio as a percentage of agreement.³³ Thus, if the ten decisions were equally divided in favour of plaintiff and defendant this equals a zero percent agreement. On the other hand, if the ten decisions were ten for one of the parties to zero for the other, this equals a one hundred percent agreement. A six to four split, to take one further example, equals twenty percent agreement. We also calculated agreement by combining singles' and pairs' decisions in each dispute.³⁴ Because the combined measure of agreement is based on twenty decisions instead of ten, the combined results may be regarded as the most reliable indications of law model effects.³⁵

33. The formula used was: % agreement = $100 \times \text{abs}(2d - n) / n$, where n is the number of decisions (10 singles or 10 pairs judgments) and d is the number of decisions in favour of plaintiff or defendant (as the absolute value of $(2d - n)$ is used, it does not matter which).

34. The calculation of agreement for the combined data uses the formula described in note 33, although the number of decisions is, of course, 20 rather than 10. Thus, 20 decisions for one party equals 100% agreement, while 10 decisions for each equals 0% agreement.

35. The measure of agreement makes the "sampling unit" the dispute, rather than the individual participant (or pair of participants), and reduces our "sample size" (number of observations in each experimental group) by a factor of 10 compared to the analysis of participant evaluations of utility. Thus each experiment yields only 30 observations (10 for each law model) where, by contrast, the number of observations of each participant evaluation of utility was 300 in Experiment 1 (singles) and 600 in Experiment 2 (pairs). Sample size affects statistical power, that is, the ability to detect significance in observed differences between law models. Although combining the singles and pairs data does not increase the sample size (each case still only yields one

When the level of agreement is calculated for all ten disputes, there is no law model effect, that is to say that no one of the law models leads to more predictable results overall.³⁶ However, in our view some of the ten disputes were easier to decide than others. We ranked the disputes in order of their difficulty, and divided them into the five "easier" and the five "harder" decisions. Figure 1 presents the levels of agreement for two groups of five "easier" and "harder" decisions. Figure 1 illustrates that when difficulty is taken into account, the law models differ significantly.³⁷ Users of the ACC agreed more often on the outcome in easier cases, and disagreed more often in harder cases. While difficulty made some difference in agreement among UPICC users, the effect was smaller. In contrast, difficulty made no difference to users of Case Law.

As previously noted, one could claim that broad principles make decisions unpredictable. The data suggest that, if anything, the reverse is true. The reader should recall that every dispute used in our experiments was based on the decision of a split appellate court. Nevertheless, in half of the disputes nearly eighteen out of twenty decisions made by law students using broad principles agreed on the result.³⁸

The study yielded several other results that indicated that broad principles make it easier to agree on the outcome,³⁹ while detailed

measure of agreement for each law model), it does increase power somewhat by reducing the variability of the individual agreement estimates, since they are based on twice as many decisions.

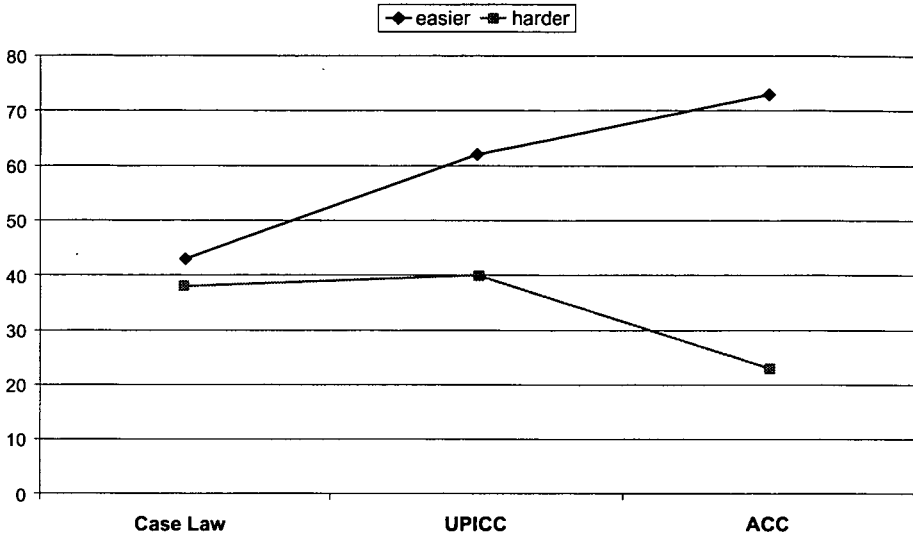
36. As is conventional, null hypothesis probabilities (p) less than .05 (corresponding to a chance of less than 1 in 20 of wrongly rejecting the null hypothesis) are described in this report as "significant"; and p equal to or greater than .05 and less than .10 as "marginally significant." A few marginally significant findings are reported below, where the data exhibits a trend or pattern of difference which is consistent with other findings. As the term indicates, however, these are less reliable than significant findings, and conclusions based on them should be regarded as provisional. In this case, although a test of the differences among the singles agreement scores using a non-parametric or rank based (Friedman) test gave $S(2) = 5.52$, $p = .063$, which is marginally significant, the test using a one-way repeated measures ANOVA gave $F(2, 18) = 2.28$, $p = .131$, which is clearly not significant. The observed differences between the pairs and combined means were clearly not significant: pairs $F(2, 18) = 0.89$, $p = .427$; combined $F(2, 18) = 0.75$, $p = .488$.

37. Law Model \times Difficulty $F(2, 16) = 4.62$, $p = .026$. The main effect of Difficulty is marginally significant: $F(1, 16) = 4.86$, $p = .059$.

38. The level of agreement among ACC users deciding the more difficult cases is approximately the same as a 2:1 split.

39. Pairs using the ACC rated as more helpful in reaching a fair result and this effect was strongest in easier cases: Law Model $F(2, 294) = 4.60$, $p = .011$; Law Model \times Difficulty $F(2, 294) = 2.95$, $p = .054$ (marginally significant). Ratings of judgments for agreement, fairness and consideration of the facts indicated that the ACC made the fair outcome more apparent, at least in easier cases. See *infra* notes 52, 54 and accompanying text. It can be assumed that if a law model makes the fair outcome more apparent to a reader of a judgment, it will also make it more apparent to a decision maker, thus increasing predictability of decisions.

Figure 1.
Mean Percent Agreement (Combined Data) by
Law Model and Difficulty



rules have a tendency to complicate even easier cases.⁴⁰ A capacity to sort out simple from difficult disputes is a highly desirable characteristic. Detailed rules seem to have a tendency to increase the scope for plausible disagreement over the outcome of simple disputes, thus possibly promoting litigation. On the other hand, broad principles appear clarify when a decision is difficult, thus possibly leading to outcomes that are more just.

B. Justice

A number of significant differences between the law models suggest that broad principles are more likely than detailed rules to yield just outcomes:

40. Pairs rating how much the rules helped them to decide, rated Case Law in easier cases lowest out of all groups: Law Model \times Difficulty $F(2, 294) = 2.95$; $p = .054$ (marginally significant). There was no difference between Case Law and UPICC ratings, and UPICC ratings did not distinguish between harder and easier cases. There was a clear trend indicating that pairs using Case Law found it harder to agree in easier cases, compared to all other groups: Law Model \times Difficulty $F(2, 294) = 2.55$, $p = .08$ (marginally significant). While UPICC pairs apparently found it easier to agree in easier cases, other data suggests that this is attributable to selective application of its provisions. See *infra* Table 1 and accompanying text; see also Ellinghaus & Wright, *supra* note 1. Readers of Case Law judgments in easier cases, compared to all other groups, found the reasons hardest to read ($F(2, 1788) = 17.13$, $p = .000$), most jargon-laden ($F(2, 1785) = 7.30$, $p = .001$), most technical see *infra* Figure 4 and accompanying text, most confusing ($F(2, 1786) = 3.10$, $p = .045$) and least easy to follow ($F(2, 1788) = 3.06$, $p = .047$).

1. Student judges applied broad principles more accurately.⁴¹
2. They rated broad principles more helpful in reaching a fair result.⁴²
3. They also rated broad principles less technically.⁴³
4. Readers of judgments based on broad principles discriminated more strongly between fair and unfair decisions.⁴⁴
5. They also rated judgments based on broad principles less technically.⁴⁵

In order to measure accuracy of application, we jointly assigned a mark out of ten to the reasons given by student judges for their decisions, based on our assessment of how accurately they applied the relevant statement of law. In doing so, we were conscious of the potential for researcher bias in this process and took a number of measures against it. We endeavoured to make the assessment criteria as objective as possible. In giving a mark, we paid no attention to the outcome. We gave marks whenever a student judge referred to a relevant rule, unless it had been plainly misunderstood. We relied on our extensive experience in marking law student answers to legal problems. The procedure implemented required one of us to read a statement of reasons aloud to the other, and then each declared when he had arrived at a mark. We took turns to be the first to announce a mark. In the overwhelming majority of cases, the secondary announcement agreed almost exactly with the primary announcement. In only a small minority was there need for discussion.⁴⁶ Table 1 shows the mean mark awarded by law model.

41. See *infra* Table 1 and accompanying text.

42. Paired users rating helpfulness rated ACC highest in easier cases, while Case Law users in easier cases found it least helpful, compared to all other groups: Law Model \times Difficulty $F(2, 294) = 2.95, p = .054$ (marginally significant).

43. Ratings in response to "The rules were technical" were significantly affected by Law Model: Singles $F(2, 292) = 3.90, p = .021$, pairs $F(2, 293) = 5.13, p = .006$. In the case of the singles, ACC judges rated it significantly less technical than Case Law, while UPICC ratings were in between. In the case of the pairs, ACC was rated significantly less technical than Case Law, and there was a strong trend towards rating the ACC better than UPICC, while UPICC and Case Law ratings did not differ.

44. See *infra* notes 50, 52 and accompanying text.

45. Readers of judgments rated ACC judgments less technical than UPICC and Case Law judgments: $F(2, 1793) = 4.06, p = .017$. The very interesting interactions between Law Model and Difficulty, and Law Model and Fairness, are discussed below. See *infra* Figures 2, 3 and accompanying text.

46. Copies of the reasons and marks assigned have been retained and are available for inspection and independent assessment. Requests should be directed to the Law & Justice Foundation of NSW.

Table 1:
Mean Mark (Out of 10) Awarded to Judgment
Reasons, by Law Model.

	Easier cases			Harder cases		
	Case Law	UPICC	ACC	Case Law	UPICC	ACC
Singles	5.6	5.0	6.8	4.8	4.5	5.7
Pairs	4.4	4.1	5.9	4.1	3.9	5.7
Combined	5.0	4.5	6.4	4.5	4.2	5.7

These results show that law students were able to apply broad principles with significantly greater accuracy than detailed rules.⁴⁷

We also analyzed the proportion of fair outcomes reached by the student judges. We defined fair outcome in two ways: as the result reached by the majority of the court in the original case, and by reference to our own opinion. We disagreed with the majority in three out of the ten cases.

We found that users of detailed rules and users of broad principles were equally likely to reach the fair outcome defined by court majority.⁴⁸ On the other hand, both users of UPICC and the ACC were more likely than users of Case Law to reach the fair outcome defined by our opinion.⁴⁹ However, this finding must be qualified. As Table 1 indicates, UPICC users did not accurately apply its provisions. It can therefore be concluded that broad principles are more likely to be accurately applied, and as a result are more likely to produce just outcomes.

Before we undertook our analysis of the quality of reasons, we suspected that UPICC users might have tended to resort directly to Article 1.7, the general duty of good faith, without reference to other, more detailed provisions. In fact, we found that the UPICC users fre-

47. Law Model $F(2, 581) = 35.321, p = .000$. The quality of reasons given for harder decisions was lower than the quality of reasons given for easier decisions: Difficulty $F(1, 581) = 9.264, p = .002$. Single judge reasons received higher marks than pair judge reasons: Single-Pair $F(1, 581) = 17.675, p = .000$. Although this result apparently runs counter to our hypothesis that pair decisions would be of better quality, it can be explained by the fact that pairs took longer to decide, and therefore had less time to write their reasons. See *infra* Table 2 and accompanying text. There are no significant interactions between these factors.

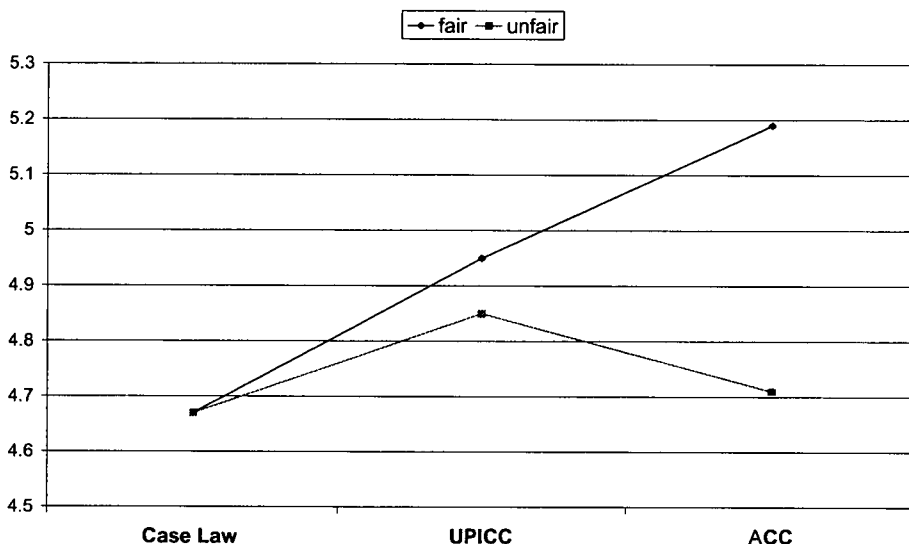
48. Law Model: singles $F(2, 16) = .06, p = .941$; pairs $F(2, 16) = .35, p = .707$; combined $F(2, 16) = 0.26, p = .777$. Law Model \times Difficulty: singles $F(2, 16) = 1.59, p = .234$; pairs $F(2, 16) = 1.59, p = .235$; combined $F(2, 16) = 2.51, p = .112$. The proportion of fair outcomes was also not affected by Difficulty: combined $F(1, 16) = 0.69, p = .432$.

49. The effect of Law Model on the singles data was not significant ($F(2, 16) = 1.99, p = .169$), although the pattern was very similar to that of the pairs data. The effect of Law Model on both pairs and combined data was significant: pairs $F(2, 16) = 3.99, p = .039$; combined $F(2, 16), p = .024$.

quently omitted any reference to Article 1.7, and showed a marked tendency to select from the detailed rules those provisions that supported the judges' decisions, and to ignore other potentially relevant rules. We suspect that selective application is a flaw to which models based on a large number of detailed rules are more susceptible than models based on a few broad principles.

The judgment readers were asked, among other things, how strongly they agreed or disagreed with the proposition that a "judgment took into account all of the important facts in this case." Figure 2 shows that readers of judgments based on broad principles differentiated more strongly between fair and unfair outcomes (researcher opinion), in easier cases.⁵⁰

Figure 2
Effect of Law Model, Fairness on Fact Content
Ratings (Easier Cases)



It is important that the law draws a decision-maker's attention to all the important facts of a dispute. This finding suggests that broad principles do this more effectively than detailed rules.⁵¹

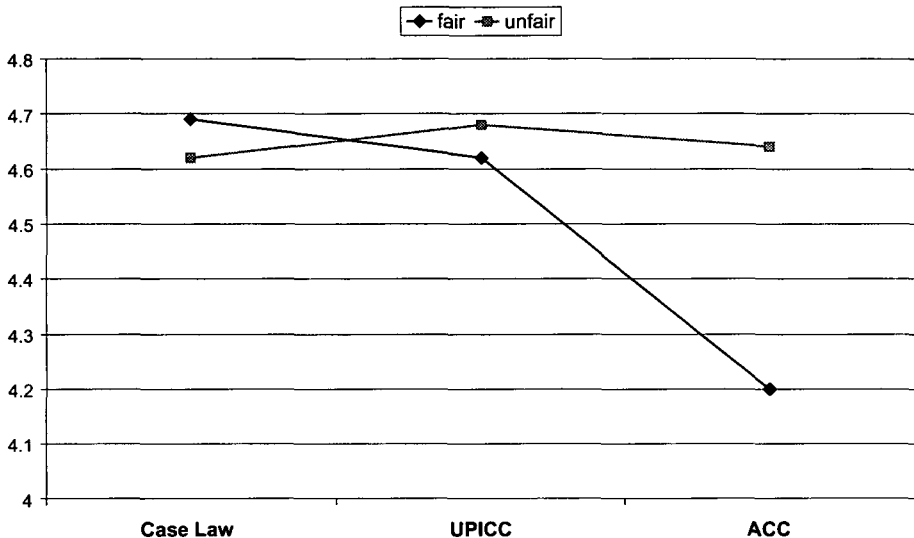
50. Law Model \times Fairness \times Difficulty $F(2, 1785) = 3.28, p = .038$. The reader must accept this conclusion as an accurate summary of a necessarily much more detailed and technical discussion of this three-way interaction contained in the full report. See generally Ellinghaus & Wright, *supra* note 1.

51. It may be thought that the Likert scale differences are small and therefore not of any importance. In other words, their ostensible numerical precision may mislead some readers. Likert scales are in fact ordinal and not interval scales. The technique does not assume that respondents regard the difference between, say, 3 and 4 as the same as the difference between 4 and 5. Indeed, it is known that respondents gener-

Readers' ratings of fairness of judgments and agreement with decisions produced very similar results. That is, broad principles enabled readers to discriminate more strongly between fair and unfair outcomes, at least in easier cases.⁵² It is an obvious shortcoming of a law model if it does not assist readers in identifying the fair outcome in easy cases. On the other hand, it is to be expected that readers will find it difficult to discriminate between fair and unfair outcomes in difficult cases.

Readers of judgments were also asked to indicate their level of agreement or disagreement with the proposition that "the decision was technical." The results are shown in Figure 3.

Figure 3
Effect of Law Model, Fairness on Technicality Ratings



The technicality ratings of readers of broad principle judgments discriminated strongly between fair and unfair outcomes,⁵³ this time de-

ally do not respond to the intervals on Likert scales as equal. Thus, while it is tempting to suggest that a 1 point difference on a 7 point Likert scale is nearly 15% (or a half point difference is 7%), strictly speaking this cannot be assumed. The technique in fact yields relative rankings of experimental groups in relation to the relevant proposition. The most, then, that one can say of any Likert scale result is whether it is statistically significant (i.e., the rankings are reliable) and, therefore, should not be dismissed as inconsequential.

52. Agreement with decision: Law Model \times Fairness \times Difficulty $F(2, 1787) = 5.21$, $p = .006$. Fairness of judgment: Law Model \times Fairness \times Difficulty $F(2, 1788) = 4.47$, $p = .012$. Again the reader must accept this conclusion as an accurate summary of a necessarily much more detailed and technical discussion of this three-way interaction contained in the full report. See generally Ellinghaus & Wright, *supra* note 1.

53. Law Model \times Fairness $F(2, 1784) = 3.98$, $p = .019$.

fined by the court majority, in both easier and harder cases.⁵⁴ It seems likely that if broad principles make the fair outcome more apparent to a reader of a judgment, then they will also make it more apparent to a decision maker, thus increasing the likelihood of just outcomes.

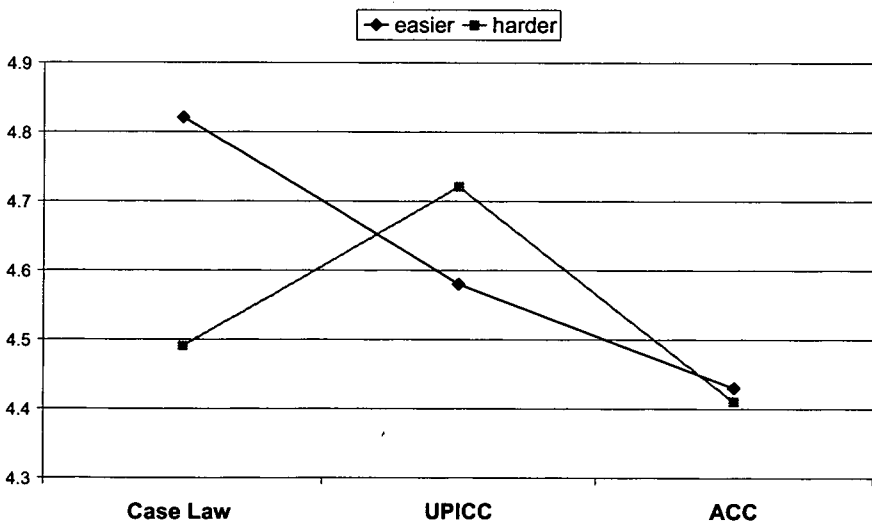
C. Accessibility

A number of significant differences among law models suggest broad principles are more accessible, use clearer language, and have clearer logic than detailed rules.

1. Student judges rated broad principles less technical.⁵⁵
2. Readers of broad principle judgments rated them less technical.⁵⁶
3. They also rated them less jargon-laden.⁵⁷
4. Detailed rules complicated easy decisions.⁵⁸

As already noted, judgment readers were asked to indicate their level of agreement or disagreement with the proposition "the decision was technical." As shown by Figure 4, judgments based on the ACC were rated significantly less technical than judgments based on Case Law or UPICC, regardless of their difficulty.⁵⁹

Figure 4
Effect of Law Model, Difficulty on Technicality Ratings



54. See *infra* Figure 4 and accompanying text.

55. See *supra* note 43.

56. See *supra* Figure 3 and accompanying text.

57. Law Model $F(21, 1794) = 43.00, p = .000$.

58. See *infra* Figure 4 and accompanying text.

59. Law Model \times Difficulty $F(2, 1784) = 3.37, p = .035$.

Another striking finding appearing in Figure 4 is that the most negative rating was given to Case Law judgments in easier cases. The reliability of this finding is reinforced by several other significant interactions of law model and difficulty, in both student judge and judgment reader data.⁶⁰ It demonstrates again the tendency of detailed rules to complicate simple disputes.

D. Efficiency

The time taken to arrive at a decision is a measure of the ease with which a law statement can be comprehended and applied. In Experiments 1 and 2, this time was measured by the interval between the start of each experimental session and the time recorded by student judges as the time they required to reach their decision. As Table 2 indicates, ACC decisions took about five minutes less than Case Law and UPICC decisions,⁶¹ a difference of about fifteen percent.⁶² Pairs took about five minutes longer than singles.⁶³ There was no interaction between the law model and the single-pair factors.⁶⁴

Table 2
Mean Time (Minutes) to Judgment by Law Model

	Case Law	UPICC	ACC
Singles	30	28	25
Pairs	35	35	29

These results suggest that broad principles are more efficient than detailed rules. As already noted, our design eliminated the need for Case Law users to locate, read, and digest relevant precedents. The fact that Case Law times are the same as the UPICC times can therefore be attributed to their greater use of detailed rules.⁶⁵

60. See *supra* note 40.

61. Law Model: singles $F(2, 294) = 2.86, p = .059$; pairs $F(2, 290) = 5.48, p = .005$; combined $F(2, 584) = 9.43, p = .000$.

62. We included the Difficulty factor in this analysis. There was an apparent trend for decisions in harder cases to take about 1 minute longer than decisions in easier cases, but this difference was not statistically significant: singles $F(1, 294) = 0.10, p = .753$; pairs $F(1, 290) = 0.45, p = .503$; combined $F(2, 584) = 2.52, p = .113$. Therefore we have not depicted Difficulty in Table 2.

63. Single - Pair $F(1, 584) = 24.57, p = .000$.

64. As is common with time data, both the ANOVA assumptions about homogeneity of variance and normality were violated. However, ANOVA is robust to these violations, which in any case were not large. In fact, the same pattern of significant results was found when these violations were reduced by logarithmic and square root transformations, and the exclusion of extreme values.

65. See *supra* Part IV.C.

VI. CONCLUSION

Let us restate our findings in broad summary. First, we found decisions applying detailed rules were no more predictable than decisions applying broad principles. However, decisions applying broad principles were significantly more predictable in easier cases. Second, we found broad principles were more likely to lead to just outcomes. Third, we found broad principles were more accessible than detailed rules. Finally, we found broad principles were significantly more efficient.

As we said in the beginning, the view that has dominated the development of the common law for at least two centuries is that detailed rules lead to more predictable and more just outcomes, and are more efficient than broad principles. Our empirical evidence indicates the reverse is true, and the hope of increasing certainty by the continuing proliferation of detailed rules is an illusion. The direct application of broad principles, which implies the relegation of mediating rules to the status of illustrations, is, if anything, more likely to produce more just and predictable results. This will not surprise some readers.

This is not the place to explore the implications of our findings for contract law or the common law in general. But, our findings suggest that more open recognition and more direct application of broad principles may well enhance the utility of case law doctrine. If we abandoned the illusion that more detail always improves the law, then this would allow some other clear advantages of broad principles, such as their potential for simplifying law, making the law more accessible to being more fully exploited.

Our research may also have transnational implications. We are now living in a global environment in which there is a growing, and probably irresistible movement, towards the international harmonisation of contract law. This objective is unlikely to be achieved except through agreement on statements of broad principles in canonical form. If the common law is to remain a world force in this new order, then it may itself have to be restated in such a form.