Coherence Based Reasoning and Models of Contract Law

Natalie Close (natalie.close@newcastle.edu.au)
School of Psychology, The University of Newcastle, Callaghan, NSW, 2300, Australia.

Andrew Heathcote (Andrew.heathcote@newcastle.edu.au)
School of Psychology, The University of Newcastle, Callaghan, NSW, 2300, Australia.

Fred Ellinghaus (f.ellinghaus@unimelb.edu.au)
School of Law, The University of Melbourne, Parkville, VIC, 3052, Australia.

Ted Wright (ted.wright@newcastle.edu.au)
School of Law, The University of Newcastle, Callaghan, NSW, 2300, Australia.

Abstract

We report the results of an experiment examining the effect of the system of rules (law model) that govern the adjudication of difficult real-world contract law disputes by senior law students. Participants rated the importance of facts that varied in relevance to the dispute throughout the course of making a decision. Similar to the coherence shift noted by other researchers studying legal decision making, the importance of facts decreased or increased consistent with their relevance as dictated by the law model and the participant's emerging decision. We simulated these results, and the interaction between the effects of dispute difficulty and law model on decision consensus, using a constraint satisfaction network.

Keywords: Legal reasoning; coherence shift; models of law.

Background

Disputes that reach adjudication within a legal system involve competing arguments of such complexity that an impartial third party is required to decide the outcome. Despite high levels of ambiguity, the presiding judge processes the facts of a case and reaches a decision that they believe is fair and due. Understanding the cognitive processes involved in making complex legal decisions is a developing area in both cognitive science and empirical legal studies. Cognitive models of legal decision making have emerged (Holyoak & Simon, 1999; Thagard, 2004; 2006) and the focus has now turned toward examining the rules that regulate complex legal decisions (Hirsch, 2006). This form of rule-based decision making can be best described as deductive reasoning, despite the analogical nature of rule creation within common law systems (see Schauer, 2008).

Models of law

Systems of legal rules or *law models* fall on a continuum between the general (*broad principles*) and specific (*detailed rules*). The nature of the best law model to govern legal decision making has been the subject of a longstanding jurisprudential debate. This debate is centred on the optimal specificity of legal rules and the influence of specificity on legal efficiency, utility and certainty. Legal certainty can be defined as consensus, which is the amount of agreement between a panel of judges who decide the same dispute. Legal scholars believe that optimal specificity is achieved when legal certainty is maximized, that is, when the rules lead decision makers to a predictable outcome.

Emerging from the specificity debate is the orthodoxy that the law is best expressed as highly specific detailed rules, commonly known as the common law, or case law. Detailed rules are believed to improve the certainty of law by guiding the user toward a predictable decision (Sullivan, 1992). The inflexible and highly specific nature of detailed rules pinpoint particular facts and evidence within a dispute that should be considered most important to a decision maker, and also renders many facts formally irrelevant. This provides adjudicators with a selective view of the facts to each dispute. Such selectivity has been characterized as making detailed rules highly efficient, encouraging decision makers to arrive at a satisfactory rule-bound decision, rather than embarking on an idealistic and exhaustive search for the best outcome (Hirsch, 2006).

However, recently an international move towards the global harmonization of private law has challenged the economic value of detailed rules. Advocates of legal codification argue that summarising law into broad principles has both economic and legal benefits (Fon & Parisi, 2004). Broad principles are less specific and thus more flexible, encouraging the decision maker to consider more facts and evidence as important to the outcome of a dispute. The broad principle model of law does not allow the decision maker a selective view of the case, instead encouraging the judge to consider a wider array of the dispute facts and evidence.

The codification movement has been supported by empirical evidence demonstrating that, contrary to orthodox opinion, broad principles guide an adjudicator's decision-making to outcomes that are no less certain than outcomes based on detailed rules. Ellinghaus and Wright (2005) provided law students with the facts from a contract law dispute and a statement of the relevant law in the form of either broad principles or detailed rules. They used ten ambiguous contract law disputes that produced split decisions by the Australian appellate court; five classified as easier to decide and five harder. Participants were asked to decide the outcome of their dispute based on the statement of law, and consensus was measured by the proportion of participants agreeing with the majority decision in each case. For the easy cases consensus was high for broad principles but only moderate for detailed rules, whereas for

harder cases consensus was moderate under both law models. These results indicate that broad principles promote greater consensus than detailed rules and that they do so to the greatest degree, as seems appropriate, in easier cases.

Models of legal decision making

Simon (1998) presented a model of coherence-based reasoning (CBR) that describes how legal decision makers turn conflicting facts into coherent arguments that support an outcome. The model assumes that over the course of coming to a decision an adjudicator unconsciously restructures the facts of a dispute into consistent inferences that support their decision. The CBR model is implemented as a connectionist constraint satisfaction network where the facts and inferences relevant to a dispute are represented by nodes connected by inhibitory and excitatory links. Excitatory links join nodes representing consistent aspects of the dispute and inhibitory links join inconsistent aspects. Node activations are updated in a way that tends to converge on a coherent representation, where a mutually consistent set of nodes becomes highly active and inconsistent nodes are suppressed (Simon, Krawczyk & Holyoak, 2004). The difference between the initial activation, which represents the original statement of facts in the dispute, and final activation, which represents the decision and supporting inferences, is called a coherence shift (Simon, Pham, Le & Holyoak, 2001; Simon, Krawczyk & Holyoak, 2004; Simon, Snow & Read, 2004; for review see Simon 1998, 2004).

Holyoak and Simon (1999) demonstrated that behaviour in a judging task was consistent with the CBR model. In particular, participants' rating of case-relevant legal inferences changed over the course of deciding the dispute. Agreement with eight legal points designed to favor either the plaintiff or defendant was moderate and undifferentiated between the parties at the beginning of the decision task (pretest). However, over the course of decision making participants organized the legal points into a coherent story that supported their eventual decision. The product of this process was a division between participants' agreement with points that support the defendant and plaintiff at the conclusion of decision making (posttest). Points that cohered with their emerging decision were rated as more important over time, while the importance of opposing points was suppressed. By comparing the pretest and posttest ratings, Holyoak and Simon concluded that, during the course of decision making, there is a separation between the initial and final mental model of the task.

Thagard (2006) argued that explanatory coherence provides a psychologically plausible account of how competing explanations are evaluated. Thagard (2004) used a computational model of causal inference (ECHO), which is related to the CBR model, to simulate a criminal law dispute. The dispute was tried twice, with the first jury finding the defendant guilty of murder, but a second jury acquitting the defendant after additional evidence was introduced. The ECHO model produced the same change of decision, demonstrating the ability of the model to represent some aspects of human decision making.

Testing Coherence in Contract Disputes

We ran an experiment to test whether the CBR framework could be used to understand the effect of the different law models on decisions in contract law disputes. In order to examine the emergence of coherence we asked participants to rate the importance of facts several times during the course of deciding a contract dispute taken from Ellinghaus and Wright (2005). To examine the effect of law model, different groups of participants were asked to use statements of the applicable laws to guide their decision making. This statement was composed of the laws expressed as detailed rules or as broad principles. We also classified a set of case facts according to their relevance to the different law models in order to determine whether any change in the ratings of these effects over the course of making a decision would be influenced by the law model being applied.

In particular, facts were classified as a) "Glue", b) "Broad Principle" or c) "Detailed Rule". Glue facts were part of the case narrative but were not relevant to the outcome of the dispute under either law model (legally irrelevant). Broad principle facts should only be relevant to participants using the broad principles law model to decide their dispute. They represent legally relevant information, but this information is dictated as irrelevant by a detailed rule law model. Detailed rule facts should be important to participants using both broad principles and detailed rules as both law models indicated they were relevant to the outcome of the dispute. If participants arrived at a decision by developing a coherent set of causal inferences we predicted that under both law models glue facts should receive increasingly lower importance ratings and detailed rule facts increasingly higher importance ratings. For broad principle facts we predicted that ratings should differ as a function of law model. Because detailed rules dictate that these facts should not be taken into account, participants should give them increasingly lower ratings. In contrast, as broad principles allow these facts to be considered, they should be increasingly rated as important by the group using broad principles to guide their decision.

Method

Participants ($N = 107$) were JD (graduate law degree) or LLB (Bachelor of Laws) students from four Australian universities who could graduate within the next two years. An invitation to participate was distributed to eligible students through the office of the Dean at each law school. Participants were reimbursed for their time with a A$40 gift voucher for amazon.com. It took approximately 1.5 hours to complete the experiment.

Materials consisted of six descriptions of contract law disputes taken from Ellinghaus and Wright (2005), three of which were classified as easier and three harder to decide. These cases embody a range of issues pertinent to contract law. The third and fourth authors (contract law experts)

extracted a total of 15 facts from each case, with equal numbers of each type (Glue, Broad Principle and Detailed Rule). They also prepared two statements of the law applying to each dispute drawing on the materials used in Ellinghaus and Wright. One statement was based on Australian Case Law (informally known as the common law). The other statement of law was based on the draft Australian Contract Code. The primary distinction between these two models of law is that Australian Case Law is comprised of numerous detailed rules, while the code contains only 27 broad principles.

The experiment was presented on the Internet and accessed remotely by participants at a time of their choosing. When they entered the site they were randomly allocated to one of the six disputes and one of the two statements of law. After reading the dispute description they rated the importance of 15 facts (5 of each type) taken from the dispute (Time 1). Ratings were made on a Likert scale with responses ranging from 1 (Not Important) to 7 (Extremely Important). On all four occasions where facts were rated the facts were presented in five groups of three, with one fact of each type in each group. The order of presentation of the facts was randomized.

Next, participants read a statement of the law they were to use in judging the dispute. During this time participants could switch back and forth between their dispute and the statement of law freely. After again rating the facts (Time 2) they were asked to assume the role of a judge and to prepare and record well-reasoned arguments for both parties to the dispute based on their law statement. The order of argument presentation was randomized between subjects. After again rating the facts (Time 3) participants recorded a decision for one of the parties and provided written reasons for that decision, based again on the rules of their statement of law. They then made a final rating of the facts (Time 4).

Results

A Generalized Linear Model analysis of the binary consensus data was performed assuming a binomial probit link function (McCullagh & Nelder, 1989). It revealed significant main effects of case difficulty ($Z=2.01$, $p=.044$) and law model ($Z=2.18$, $p=.03$). Simple effects analysis showed that there was a significantly greater consensus for broad principles than detailed rules in easy cases ($Z=2.18$, $p=.03$) but no reliable effect of law model for hard cases ($Z=0.56$, $p=0.57$).

Figure 1 shows mean importance ratings for each fact type throughout the experiment for each law model group[1]. There were highly significant interactions between fact type and time, $F(6,618)=5.4$, $p<.001$, and between fact type, time

---

[1]For clarity, Figure 1 averages over case difficulty as difficulty did not participate in any reliable effects except the four-way interaction among all factors, $F(6,618)=2.52$, $p=.02$. Simple effect analyses of each fact type showed this interaction was due to ratings of glue facts. In particular, ratings for the two law model groups in hard and easy cases differed as a function of time in an apparently unsystematic manner, $F(3,309)=3.11$, $p=.027$.

and law model, $F(6,618)=2.95$, $p<.01$. Simple effects analysis of each fact type showed that these interactions were due to a) increasing ratings with time for detailed rule facts, $F(3,309)=3.26$, $p=.022$, b) decreasing ratings with time for glue facts, $F(3,309)=4.98$, $p<.01$, and c) a marginally significant interaction between the effects of time and law model for broad principal facts due to an increase with time for the broad principal law model group a decrease with time for the detailed rules law model group, $F(3,309)=2.56$, $p=0.055$.

Table 1: Consensus with the majority (%, corresponding frequencies given in brackets) for participants who used either detailed rules or broad principles to judge disputes classified as easier or harder to decide.

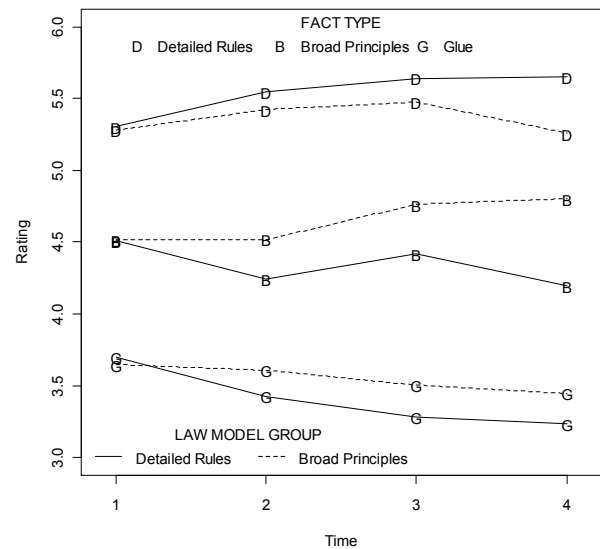|	Broad Principles	Detailed Rules
Easier Disputes	89% (25/28)	64% (16/25)
Harder Disputes	67% (18/27)	59% (16/27)

Figure 1: Rating of fact importance averaged over all disputes as a function of time (1=after reading the dispute, 2=after reading the law model, 3=after constructing arguments and 4=after making a decision), for different fact types and groups using different law models.

Discussion

The results of the experiment indicate that the explanatory coherence framework can be used to understand the effect of the different law models on decisions in difficult contract law disputes. Holyoak and Simon (1999) investigation of coherence shifts in a legal context examined how agreement with pro-plaintiff and pro-defendant inferences changed over the course of coming to a decision. They found that ratings of agreement with inferences congruent to the eventual decision increased, whereas agreement decreased with inconsistent inferences. We used a similar design but examined ratings of the importance of different types of

case facts. We found a phenomenon related to the coherence shift reported by Holyoak and Simon whereby the relevance of facts to the eventual decision determined how ratings changed over the course of making a decision. Our participants rated facts that were highly relevant to the inferences that supported their eventual decision (e.g., Detailed Rule facts) as increasingly important. Conversely, facts that were irrelevant to the decision (e.g., Glue facts) were rated as increasingly unimportant.

Our design also allowed us to investigate how the set of rules used in adjudication (the law model) affects the adjudication process. We identified a set of facts (Broad Principle facts) that were relevant to the decision under one law model (broad principles) but which are classified as irrelevant to the decision under the other law model (detailed rules). Before being introduced to their law model, participants rated detailed rule facts as most important and glue facts as least important with broad principle facts falling in between. Over the course of coming to a decision these initial differences increased in approximately the same way under both law model for the most and least important (i.e., detailed rule and glue) classes of facts. However, the change for broad principle facts depended on the law model. When participants based their adjudication on detailed rules, which classify broad principle facts as irrelevant, their importance ratings decreased. Conversely, broad principle facts were classified as increasingly important by participants who based their adjudication on broad principles.

In summary, ratings of fact importance, like ratings of agreement with legal inferences, change over the course of decision making. These changes depend on the relevance of the facts to making a decision, with more relevant facts receiving increased ratings and less relevant facts decreased ratings. Two factors determined fact relevance in our experiment. First, after reading a dispute, but before being given instruction as to the rules by which a decision should be reached, participants strongly differentiated between the importance of the three classes of facts which we examined. Presumably such differences reflect participants apriori knowledge of law (as our participants were senior law students). The second factor was the explicit law model by which they were asked to judge case. Facts that were rated equally important before seeing the law model either increased or decreased in importance depending on whether the law model classified them as relevant or irrelevant. Although reliable, this effect was small relative to the initial differences between fact types, indicating that participants' prior beliefs had a strong effect despite instructions to participants to base their decisions on their statement of law.

Our results replicated Ellinghaus and Wright (2005) in that there was greater consensus amongst participants basing their decision on broad principles than among participants basing their decisions on detailed rules, mainly due to greater consensus in easier disputes (see Table 1). This phenomenon suggests that easier disputes more clearly favored one or other party relative to harder disputes,

mainly because of facts that would be considered relevant under broad principles but irrelevant under detailed rules (i.e., Broad Principle facts). When adjudication is guided by detailed rules the influence of broad principle facts is reduced, and so the difference between easy and hard disputes is also reduced. This effect was a very strong one, with the difference in the proportion of participants agreeing with the majority between easier and harder disputes being more than four times greater under broad principles than detailed rules. This strong effect contrasts with the effects of these factors on fact ratings; a weak but reliable effect of law model and a null effect of hard vs. easy cases.

In the next section we develop a constraint satisfaction network model of our task influenced by Simon (1998) and Thagard (2004). The model represents a preliminary exploration of how the effects of rules which guide adjudication can be integrated into the CBR framework. At this preliminary stage, our focus was on determining whether such a model could show the qualitative pattern of effects seen in our experiment rather than providing a quantitative account.

Integrating Law Models & Coherence Networks

Our model development was guided by a desire to implement the details of our experiment in a simple manner. To this end we represented the effects of different types of facts (e.g., Glue, Broad Principle and Detailed Rule), levels of evidence (e.g., evidence in harder and easier disputes) and law models (broad principles and detailed rules) through units which provide constant inputs to a constraint satisfaction network. Likely a dynamic representation of these factors (e.g., as units with changing activation) might provide as good, if not a better account, but we do not explore that possibility here.

We also sought an abstract and general characterization of our model by simulating a large number of networks that captured the variety among participants and cases. To that end we used probabilistic processes to determine connection weights for each network, rather than hand coding connections to represent relationships specific to a particular case (e.g., Thagard, 2004). Each random network can be thought of as a particular dispute-participant combination drawn from a population. Once again, this does not deny the importance of Thagard's more veridical approach. To facilitate such developments in future work we assumed binary (±1) valued connection weights, which make it simpler to code such dispute specific relationships.

The constraint satisfaction network consists of two pools of units representing causal inferences favoring the two parties to the dispute. Our simulations used 15 units of each type, but the same pattern of results was found for smaller or larger numbers. These defendant and plaintiff "story units" are connected in the standard way for constraint satisfaction networks: symmetric excitatory connections within a pool (made with probability p_{dd} and p_{pp}), symmetric inhibitory between pools (made with probability p_{dp}), and no self-connections. Unit activations are initially set at a

constant low value (we used 0.01) and on each iteration a story unit is randomly chosen and its activation updated. Each update can be considered analogous to a participant considering and adjusting their assessment of the plausibility of an inference contingent on the state of other inferences and sources of evidence. Assuming unit i has a total input:

$$net_i = \sum_j A_j W_{ji}$$

The summation is over units with activations A_j that have connection weights, W_{ji}, to unit i. The update on iteration t is:

$$A_i(t+1) = \begin{array}{ll} A_i(t) \times d + (1-d)(1-A_i(t)) \times net_i & net_i \geq 0 \\ A_i(t) \times d + (1-d)(1+A_i(t)) \times net_i & net_i < 0 \end{array}$$

This update equation bounds activation between ± 1 for d close to one, which avoids undershoots and overshoots of these bounds by this discrete update scheme. We used d=0.99, but found the same results for similar values.

After repeated updates the network will converge so that further updates do not cause any more change in activation. At any point a decision can be read off by determining which pool has the greatest summed activation (i.e., has the most excited and least inhibited units overall). Depending on the relative strengths of excitation and inhibition both defendant and plaintiff units might increase to a steady positive value (weak inhibition, e.g., both parties have stories that do not contradict each other) or one pool may become all positive and the other negative (strong inhibition, one story is more plausible and contradicts the other) or something in between (e.g., some for one party are suppressed even though that party wins, whereas others remain plausible for the losing party). Our simulations used equal probabilities of forming excitatory (p_{dd} and p_{pp}) and inhibitory (p_{dp}) connections and a low value (0.1). The absolute of level of this probability did not greatly affect the pattern of results but changes in the relative levels of excitation and inhibition had strong effects.

We modeled performance before the introduction of a law model by assuming only inputs from "fact" units. These units have constant activations of one and excitatory connections to story units. Units representing the three types of facts were assumed to have a lower (Glue, p_g), medium (Broad Principle, p_b) and higher (Detailed Rule, p_d) probability of being connected to story units. This reflects the apriori probability that a participant believes a fact type is related to a causal inference relevant to the dispute. Our simulations used 15 fact units, with results being similar for smaller and larger numbers. We also assumed equal numbers of units of each type; differences in numbers would have a similar effect to differences in probabilities. The probability of deciding for either party, and hence dispute difficulty, is determined by the relative magnitudes of fact connection probabilities for defendant p_d=(p_{gd},p_{bd},p_{dd}) and plaintiff p_p=(p_{gp},p_{bp},p_{dp}) units. Larger probabilities for a pool increase the chance that it will develop higher activation and hence be chosen as victor. The strength of a facts was measured by the sum of the activations of the story units to

which it connected (i.e., for fact unit j: $S_j = \Sigma_i A_i W_{ji}$, where A_i is story activation and W_{ji} is a fact-story weight). Importance for a class of facts was assumed to be proportional to the sum of strengths of units within a class divided by the sum of all strengths.

Rules were also represented by units which, like fact units, had a constant activation of one. However, in contrast to fact units, rule units could have excitatory or inhibitory connections to story units. The probability of an excitatory connection (r) differed depending on the law model being simulated, with a connection always being set to inhibitory if it was not excitatory. For the detailed rules model this probability for story unit i was set to r_i=D/(D+B+G), where D, B and G are the number of connections which the story unit receives from Detail Rule, Broad Principle and Glue fact units. For example, if a story unit received only connections from detailed rule facts (i.e., it was a pure detailed rule inference) it must be excited by the rule units, whereas if it received no such connections it must be inhibited. For the broad principles model r_i=(D+B)/(D+B+G), so, for example, a story unit would be inhibited if it was connected only to glue facts. As a consequence of these assumptions the detailed rule units tended to have a greater inhibitory effect than broad principle rule units and that inhibition particularly differed as a function of inputs from broad principle facts (e.g., a story unit with only broad principle fact inputs must be inhibited under the detailed rules model but excited under the broad principles model).

Figure 2 illustrates average results for 10000 networks of each type under settings that simulated easier and harder disputes that favored the defendant. For the harder disputes glue, broad principle and detailed rules to defendant story unit connections were generated with p_d=(.15,.65,.65) and for plaintiff story units with p_p=c(.15,.15,.55). Note that both broad principle and detailed rule facts favor the defendant, but to a much greater degree for the broad principle facts. For the easy cases p_d=(.15,.43,.5) and for plaintiff story units with p_p=c(.15,.32,.45) and so the balance of evidence favoring the defendant is weaker, although it is still greater for broad principle than detailed rule facts. Each network was first run though 10 updates with only fact inputs, with the resulting fact importance values being the Time 1 points in Figure 2. The network was then run through a further 40 updates with importance being read off after updates 15, 25 and 50 (by 50 cycles importance was no longer changing although story unit activation did change further after this). Figure 2 also gives the percentage of networks that had higher total activation in the defendant than plaintiff story units pools (i.e., that decided for the defendant and thus exhibited agreement with the majority) after 50 updates. Because of the large number of networks simulated the results in Figure 2 are quite accurate and so there was no need to perform and statistical tests on these results.

Figure 2 shows a pattern of results that is quite similar in some aspects to the results of the experiment given in Table

1 and Figure 1. Agreement with the majority was greater for easy than hard cases under both law models but this effect was more than four times greater for hard than easy cases. Despite this large effect on decisions the pattern of fact importance results was quite similar for hard an easy cases.

Detailed Rules: 60 % Broad Principles: 65 %

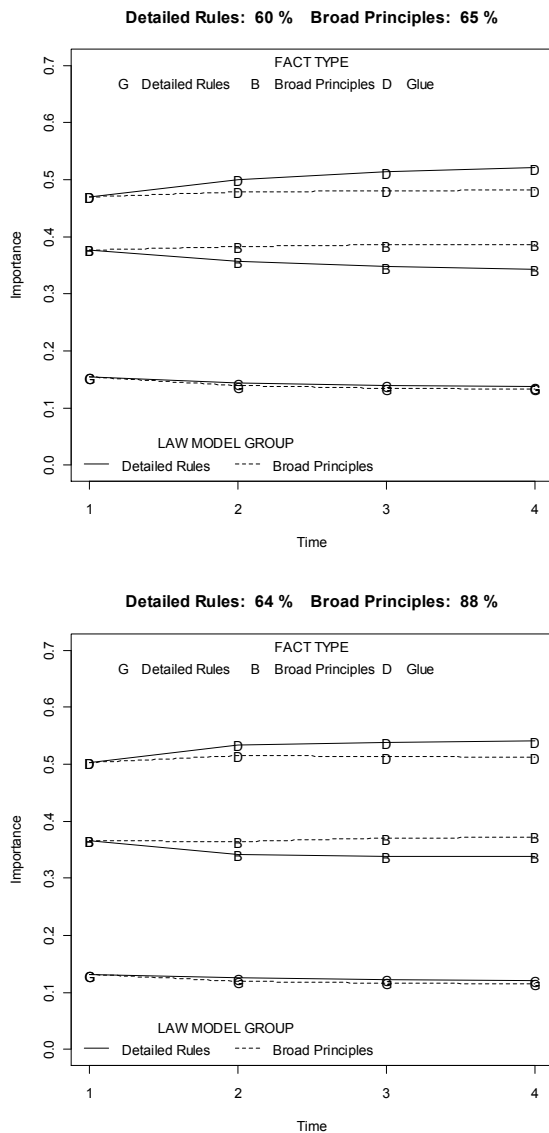**Detailed Rules: 64 % Broad Principles: 88 %**

Figure 2. Fact importance and percentage of pro-defendant decisions for easier (top panel) and harder (bottom panel) cases for detailed rule or broad principle networks.

In moving from the network with only fact inputs (Time 1) to both fact and rule inputs (Time 2) detailed rule facts were rated as more important and glue facts less important, although the latter effect was quite small. As in the empirical data these effects continued but at a decreasing rate, for later times although for the network glue importance was less differentiated between law models than in the data. The change importance with time for broad principle facts depended on law model with a clear decrease

under detailed rules and a small increase for broad principles. However, the difference between law models was smaller than in the data.

Discussion

The results of the simulation are in reasonable agreement with the data from our experiment given only four parameters (p_{bd}, p_{dd}, p_{bp}, p_{dp}) were varied to model the effects. Likely a much closer fit could have been obtained by a search of this parameter space, and by varying other parameters. For example, we observed larger changes in importance with time when inhibition was increased, either within the story units or by simulating more rule units. In any case, the present results seem to confirm the usefulness of coherence based reasoning models as providing a framework for understanding the effects of law models.

References

Ellinghaus, M.P., & Wright, E.W. (2005). *Models of Contract Law: An empirical evaluation of their utility*. Sydney: Themis Press.

Fon, V., and Parisi, F. (2004). Codifications and the optimal specificity of legal rules. *George Mason Law and Economics Research Paper, 4*(32).

Hirsch, A. J. (2006). Cognitive jurisprudence. *Southern California Law Review, 76*, 1331-1370.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. Journal of Experimental Psychology: General, 128, 3-31.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295–355.

McCullagh, P. & Nelder, J.A. (1989). *Generalized Linear Models*, Chapman & Hall.

Schauer, F. (2008). Why precedent in law (and elsewhere) is not totally (or even substantially) about analogy. *Perspectives on Psychological Science,* (6), 454-460.

Simon, D. (1998). A psychological model of judicial decision making. *Rutgers Law Journal, 30,* 1–142.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science.15*(5), 331-336.

Simon, D., Pham, L. B., Le, Q. A., & Holyoak, K. J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 1250-1260.

Simon, D., Snow, C. J., & Read, S. J. (2004). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology, 86*(6), 814-837.

Sullivan, K. M. (1992). The Justices of Rules and Standards. *Harvard Law Review106* (22).

Thagard, P. (2004). Causal inference in legal decision making: Explanatory coherence vs. Bayesian networks. *Applied Artificial Intelligence*, 18, 231-249.

Thagard, P. (2006). Evaluating explanations in science, law, and everyday life. *Current Directions in Psychological Science*, 15, 141-145.