User interaction with legal knowledge-based systems

Jaap J. Dijkstra
Section Law and IT, Department of Legal Theory,
Faculty of Law, University of Groningen,
PO box 716, 9700 AS Groningen, the Netherlands,
j.j.dijkstra@rechten.rug.nl

Abstract. This paper gives an overview of my PhD study into the persuasiveness of (legal) knowledge-based systems[1]. The results of three experiments show the possible problems that may arise when computerised legal decision aids are put into practice. The users in the experiments had great difficulties with judging the limitations of the legal knowledge-based systems they were using. For the average lawyer the competent use of computerised legal decision aids may prove to be more difficult than solving legal problems without the help of a computer. These results underline the importance of studying the usage of legal knowledge-based system.

1. Introduction: user problems

There are several studies concerning the possible improvement of decision making with the help of computer programs. Some of these studies show that the use of computerised decision aids can improve users' confidence in decision making without correspondingly improving the quality of their decisions. An example of such a study is an experiment by Aldag and Power [2]. In this experiment subjects made a strategic-management decision with the help of a computerised decision-analysis aid. They also made a decision without the help of the decision aid. Independent raters reviewed the subjects' decision reports. The results showed that the computerised decision-analysis aid did not improve the quality of the subjects' decision reports. However, when subjects used the computerised decision-analysis aid, they exhibited more confidence in the decisions they made.

In an experiment on a financial statements analysis task, Davis, Lohse, and Kottemann [3] also found that subjects misjudged the helpfulness of information systems. In their experiment an information system provided subjects with redundant (group 1) or non-redundant (group 2) information about a financial case. Compared to the condition in which subjects did not use an information system (group 3), non-redundant as well as redundant information reduced forecast accuracy while increasing confidence. These results confirm the findings of Aldag and Power: subjects had faith in an information system that did not improve their decisions. Moreover, this study shows that subjects put confidence in the information system, whatever the content of the information.

Davis and Kottemann [4] studied the 'illusion of control' effect of a method called what-if aid. What-if aid is a method for manipulating business models by specifying alternative values of decision variables and environmental assumptions. The computer program solves the model and then displays the predicted results. Davis and Kottemann

[1] This paper is based on unpublished parts of my PhD thesis [1].

Jaap J. Dijkstra, 'User interaction with legal knowledge-based systems' in: Joost Breuker, Ronald Leenes and Radboud Winkels (eds.), *Legal Knowledge and Information Systems. Jurix 2000: The Thirteenth Annual Conference*. Amsterdam: IOS Press, 2000, pp. 11-21.

found that subjects supported by a what-if analysis did not fulfil a decision task better than unaided decision-makers. Yet, the users believed that the what-if aid was beneficial and they overestimated how much it helped to improve their performance. In an additional experiment subjects believed that decision making with the help of what-if analysis was superior to unaided decision making, but they actually performed significantly worse when they used the what-if aid [5]. In a literature review of experiments on user preference versus performance Nielsen and Levy [6] found that there are many cases in which users are satisfied with computer systems that actually make them perform worse.

All in all, these experiments show that users sometimes have an unjustified confidence in decisions based on computerised information. In this paper I want to look more in detail at unjustified confidence in legal knowledge-based systems (LKBSs). I also want to give an overview of social scientific models that might explain this phenomenon. The Elaboration Likelihood Model and the Heuristic-Systematic Model focus on persuasiveness of advice. The Technology Acceptance Model tries to explain why people want to use a specific computer program.

2. Models of persuasion and technology acceptance

The Elaboration Likelihood Model (ELM [7,8]) is a social psychological model concerning persuasion. Persuasion is the psychology of attitude change. The term elaboration in the ELM refers to the extent to which people think about issue-relevant arguments contained in a persuasive message.

When situational and individual differences ensure high motivation and ability for issue-relevant thinking, the elaboration likelihood is said to be high. Consequently, the probability is relatively high that recipients of the message follow the central route to persuasion (see Figure 1). Persuasion through the central route is likely to result from a person's careful examination of the information contained in the message.

When the motivation and ability are low, the probability is relatively high that recipients of the message will follow the peripheral route to persuasion (see Figure 1). Persuasion through the peripheral route means that attitudes are determined by positive or negative cues in the persuasion context that either become directly associated with the message position or permit a simple inference about the validity of the message. The term peripheral cue refers to any variable that can affect persuasion without affecting argument scrutiny. For instance, a scientific report presented by someone who looks like a professor can be more convincing than the same report presented by someone who looks like a cheerleader.

To summarise, the ELM points out that people may adopt attitudes on bases other than their understanding and evaluation of a persuasive message. The ELM has been tested and confirmed in many experiments and it has contributed to a resurgence of interest in persuasion processes among social psychologists.

The Heuristic-Systematic Model [9] is also a social psychological model concerning persuasion. In the Heuristic-Systematic Model a distinction is made between heuristic information processing and systematic information processing. Heuristic processing is a more limited mode of information processing and it requires less cognitive effort and fewer cognitive resources than systematic processing. When processing heuristically, people focus on that subset of available information that enables them to use simple decision rules to evaluate a message. For example, the recipient of a message may look only at the message source to evaluate a message, using the heuristic that the advice of an expert can be trusted. This type of processing is similar to the peripheral route in the ELM.

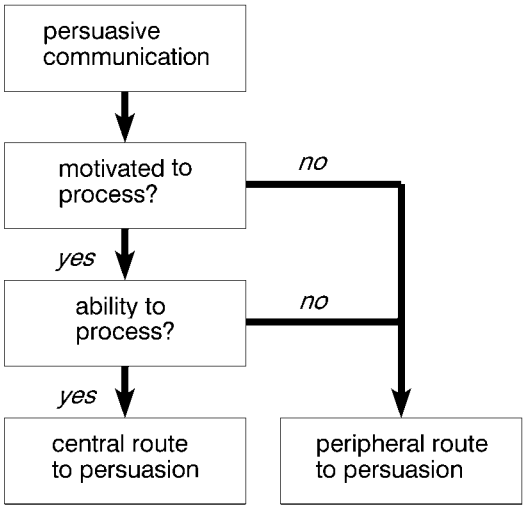


Figure 1: A simplified version of the Elaboration Likelihood Model (ELM).

Systematic processing is similar to the central route in the ELM [10]. When processing systematically, people focus on their understanding and cognitive elaboration of the argumentation contained in the persuasive message.

To explain which type of processing people use to evaluate a message, Eagly and Chaiken [9] assume that people are 'economy minded souls' who wish to satisfy their goal-related needs in the most efficient ways possible. Their least effort principle says that people prefer less effortful to more effortful information processing. People try to be as efficient as possible in their information processing. The Heuristic-Systematic Model embodies the idea that efficient information processors must strike balance between satisfying motivational concerns and reducing their processing efforts. People will exert whatever effort is required to attain a 'sufficient' degree of confidence that they have satisfactorily accomplished their processing goals.

The Technology Acceptance Model (TAM) is a model that predicts the use of information systems [11]. TAM explains the behaviour of users of information systems by looking at the beliefs these users have towards the information system. In TAM, computer use is thought to be affected by the behavioural intention that is influenced by the attitude of the user towards using a computer program. Davis [12] specified two beliefs that affect attitudes towards computers (see Figure 2). The most important belief is labelled perceived usefulness, which is the user's belief whether a computer system will be of help to attain his or her goals. The other belief is labelled perceived ease of use, which is the user's belief whether use of a computer system will be relatively free of effort.

The HCI research community has always heavily emphasised ease of use in design (user friendliness). However, results from the TAM research show that lack of user friendliness is not the most important barrier to user acceptance [11]. Users may be willing to tolerate a difficult user interface in order to use an information system they think is beneficial, while no amount of ease of use can compensate for a system that does not do a useful task.

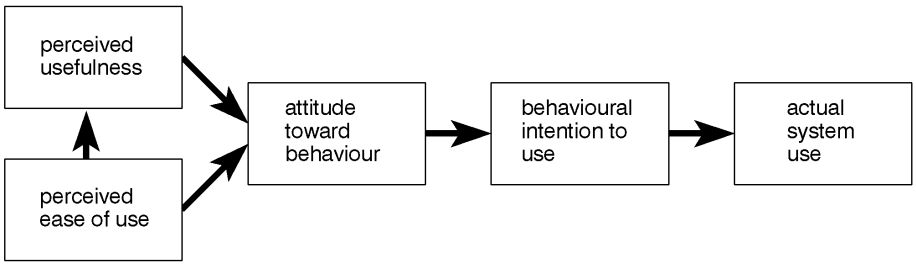


Figure 2: A simplified version of the Technology Acceptance Model (TAM)

Davis [12] refers to the work of Payne and colleagues (see [13] for an overview) in his theoretical foundation of TAM. Their cost-benefit framework focuses on trade-offs between quality and effort in decision making. This is very similar to ELM and the Heuristic-Systematic Model. Todd and Benbasat [14, 15, 16] have studied the cost-benefit framework in relationship with the use of Decision Support Systems. They found that Decision Support System use can lead to more efficient decision making with no impact on quality. Maybe decision-makers place higher value on reducing effort than on attaining the best possible decision quality, because feedback on effort expenditure is relatively immediate while feedback on accuracy is subject to both delay and ambiguity. Duan, Edwards, and Robins [17] report that in their experiments users said that reduction of effort and time was their main reason to use a computerised decision aid (but in fact the decision aid users spend more time on the task). In their study of the use of information systems Hill, Smith, and Mann [18] also mention the cognitive laziness of the user.

Research on TAM has shown that perceived usefulness is an important factor for the user's motivation to interact with a computerised decision aid. However, the user's perceived usefulness of a computer program may lie in the amount of cognitive effort that can be saved. In other words, maybe users rather prefer the computer to make the decision than to think about the decision or the applicability of the computerised advice themselves. When such a computer program is a persuasive messenger, then the user will most likely uncritically adopt the advice. This is, in a nutshell, the main topic of my PhD research [1].

Next, I want to present three experiments on persuasiveness of legal knowledge-based systems (LKBSs). Actually, the LKBS used in the first and the third experiment is not a real LKBS but a simulation designed for laboratory experiments. However, for the user it worked like a real rule-based backward reasoning LKBS (see Appendix B). The only difference with a real LKBS was that part of the input came from a database and could not be altered. This was done for experimental control reasons.

3. The first experiment: how to fool a lawyer

In the first experiment [19] 33 law school students and 30 law school teachers were asked to judge nine legal cases concerning criminal attempt (see Appendix A for an example). The lawyers were supported by one of two rule-based LKBSs that evaluated the cases and gave advice. However, both LKBSs had incomplete knowledge models. This made a good evaluation of the cases impossible because not all the important aspects of the legal cases were taken into consideration. The first LKBS deemed the suspect guilty if the acts of the suspect were dangerous for the community. This is known as the objectivistic theory.

The second LKBS focused on the intentions of the suspect to commit an offence (see Appendix B). This is known as the subjectivistic theory.

When these criminal attempt theories are used in a rigid way, they can lead to different conclusions. For example, when you consider the intentions of a pickpocket, the

attempt to steal from an empty pocket is punishable. But, when you only take the act into account, then the attempt is not punishable. Trying to take something that is not there is not endangering society and therefore such an attempt is not criminal. Currently most lawyers use a compound criminal attempt theory. Although the objectivistic elements are more important, subjectivistic elements have to be considered when judging a criminal attempt case. Thus, when lawyers judge a criminal attempt case, they have to consider both the suspect's acts and intentions.

Both the objectivistic and subjectivistic LKBS had an incomplete knowledge model (see Figure 3) and as a result they gave opposite conclusions for each of the nine legal cases in the experiment.

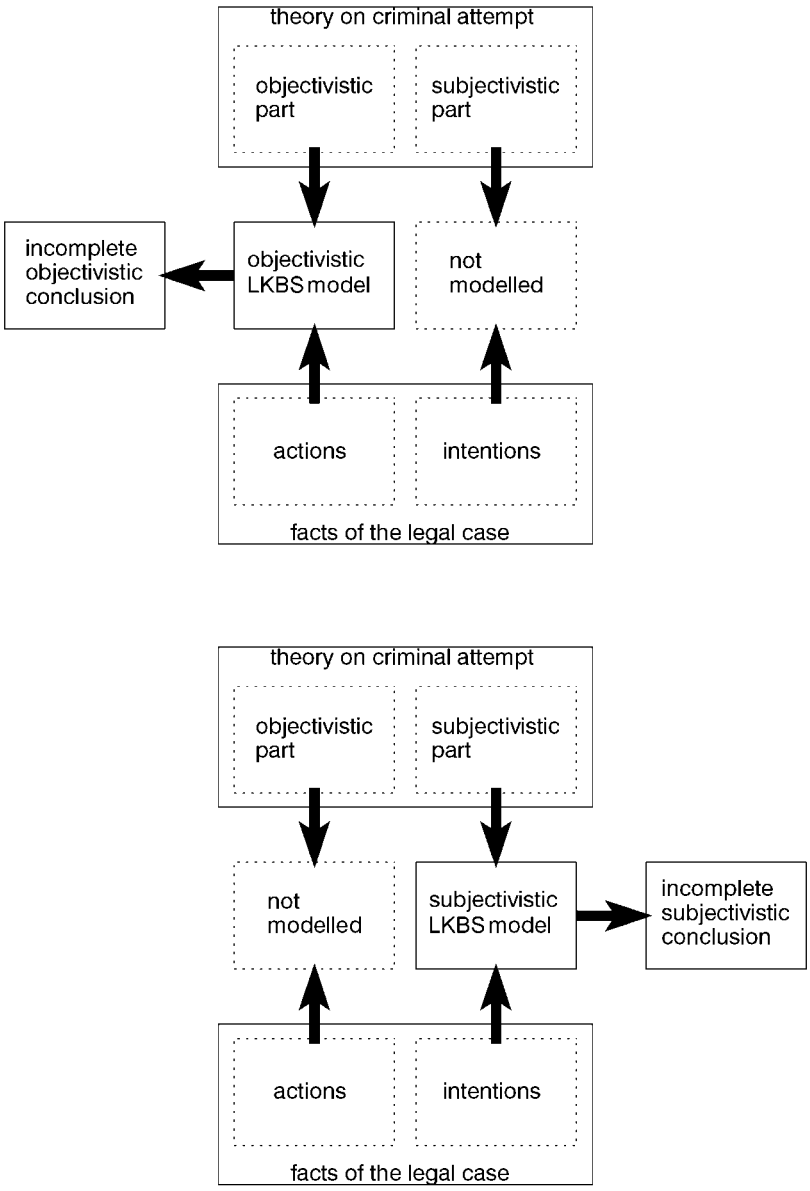


Figure 3: The incompleteness of the objectivistic LKBS (top) and subjectivistic LKBS (bottom).

The objectivistic system always deemed the suspect not-guilty and the subjectivistic system always deemed the suspect guilty. If the lawyers, who all had studied criminal attempt in law school, would critically examine the content of their LKBS argumentation, then they

would notice that the advices they were getting were based on an incomplete knowledge model. However, if the lawyers would be persuaded through the peripheral route by the advice of their supporting system, they would answer in line with their LKBS.

The results of the experiment showed a clear relationship between the LKBS used and the number of times the lawyer concluded guilty. A one-way T-test showed a significant difference between the objectivistic group and the subjectivistic group. With the teachers, the subjectivistic LKBS group scored significantly more times guilty ($M = 6.33$) than the objectivistic LKBS group did ($M = 1.33$), $t(28) = 6.60$, $p<.01$. With the students, the subjectivistic LKBS group also scored significantly more times guilty ($M = 5.63$) than the objectivistic LKBS group did ($M = 3.00$), $t(31) = 3.38$, $p<.01$. When the scores of the teachers and students were combined, the difference between the subjectivistic LKBS group ($M = 5.97$) and the objectivistic LKBS group ($M = 2.22$) was also significant $t(61) = 6.70$.

In line with the ELM and the least effort principle, the subjects hardly used the explanation functions ('how', 'why' and 'the original law text') of the LKBS to examine the advice. Only few subjects consulted an explanation function after the first forced usage in an instructional case. These results suggest that an LKBS can be a persuasive messenger.

4. The second experiment: the objective computer

The second experiment [20] tried to explain what makes (L)KBS users put their trust into computerised advice. In the experiment 85 students were asked to evaluate advices on several (legal) cases by filling out a questionnaire about the advice. The advice was given on paper. Some subjects were told that a computer ('an expert system') had made the advice and the others were told that a human expert had made the advice. A third group was also told that a computer made the advice, but this group got the advice in a simple production rule style in stead of natural language. All groups studied the same four cases and filled out the 23-item questionnaire after each case.

With a Simultaneous Component Analysis [21] the 23 items were summarised into four factors. The factors were labelled: compliance with the advice, objectivity/rationality of the advice, easiness of the case and authority of the advice (see [20] for the experimental material, the questionnaire and the scale construction).

The results of the experiment showed that subjects thought that the 'computerised advice' was more objective and rational than the same 'human advice' ($F(2,75) = 6.01$, $p < .01$), especially when the advice was given in a production rule style. Thus, perceived objectivity and rationality seems to be a persuasive belief people have about computerised advice that can make them accept the advice when they follow the peripheral ELM route. Compliance and authority did not significantly differ between the three experimental groups.

Subjects also thought that the cases were easier in the production rule style group ($F(2,75) = 6.74$, $p < .01$). An explanation for this effect could be that these subjects put an effort in trying to understand the production rule argumentation and neglected the difficulties of the cases involved.

In a similar experiment, Murphy and Yetmar [22] report that supervisors took more confidence in reports believed to be made by a colleague who used an expert system than the same reports that were 'not made with the aid of an expert system'. The supervisors also agreed more often with the 'expert system reports' and they judged the answers they perceived as computer answers as being trustworthier and more comprehensible.

5. The third experiment: cognitive laziness

The third experiment [23] is a combination of the previous two experiments. In this experiment 71 subjects used the subjectivistic LKBS from the first experiment to get computerised advice on three criminal attempt cases (see Appendix A and B). For 36 subjects the LKBS presented its inferences in natural language; the other 35 subjects got the LKBS inferences in a production rule style. Because of its incomplete knowledge model (see Section 3 and Figure 3), the LKBS always gave incorrect advice. For each case the correct advice was given to the subjects on paper.

After each case the subjects filled out the questionnaire from the second experiment and the NASA Task Load Index (NASA-TLX [24]) which is a subjective mental effort measurement. After completing the experiment the subjects filled out the TAM questionnaire on perceived usefulness and ease of use of the LKBS and a small recall test about the information presented in the experiment.

Despite the correct advice on paper, the subjects agreed in 79% of their answers with the incorrect advice of the LKBS. As could be predicted from the first experiment, the LKBS was very persuasive. For further analysis a contrast was made between subjects who always agreed with the LKBS (n=36) and subjects who disagreed once or more with the LKBS (n=35). The analyses below are 2 (natural language / production rule) by 2 (agreement / disagreement) by 3 (case) repeated measurement MANOVA tests.

Subjects who always agreed with the LKBS scored higher on the compliance scale ($F(1,67) = 34.74$, $p < .001$). All groups rated the LKBS high on the objectivity/rationality scale; the average score resembled the ‘agree’ option on the Likert scale. However, there was an interaction effect between agreement and case difference. In one case, the eggshell case, the LKBS deemed a suspect guilty for attempt of manslaughter for throwing a slipper at his wife. This man wrongly assumed that his wife had a so-called eggshell skull. The subjects who disagreed with the LKBS scored significantly lower on objectivity and rationality for this case ($F(2,134) = 4.56$, $p < .05$), but on average they still thought the advice to be objective and rational.

Subjects who always agreed with the LKBS thought the LKBS to be more authoritative than those who disagreed once or more ($F(1,67) = 8.21$, $p < .01$) and these subjects also rated the cases as being easier ($F(1,67) = 9.30$, $p < .01$).

Very much in line with the ELM and the Heuristic-Systematic Model, agreement clearly influenced the mental effort scores. Subjects who disagreed with the LKBS put more mental effort in their task ($F(1,67) = 8.92$, $p < .01$) especially those who used the LKBS with production rule style ($F(1,67) = 3.90$, $p < .05$). Disagreeing subjects also performed better on the recall test ($t(69) = 1.71$, $p < .05$).

Analysis of the use of the explanation function of the LKBS only suggests that the subjects who used the production rule style LKBS used the text function more often to get a natural language explanation about the LKBS advice (see Appendix B).

A 2 (natural language / production rule) by 2 (agreement / disagreement) ANOVA showed that the LKBS with production rule style was rated as being less useful ($F(1,67) = 6.10$, $p < .05$) and more difficult to use ($F(1,67) = 5.76$, $p < .05$). Subjects who always agreed with the LKBS rated the LKBS as being more useful ($F(1,67) = 8.72$, $p < .01$).

Summarising the results, the experiment shows that subjects who disagreed with the advice of the LKBS experienced more mental effort, had a better recall of the test material, and evaluated the cases as being more difficult than the subjects who agreed with the LKBS did. Thus, subjects who disagreed seem to have studied the advice more carefully than subjects who agreed with the LKBS did. This shows that the rejection of the advice is associated with the central route to persuasion; the evaluation of the advice is based upon a

thorough and effortful examination. The acceptance of the advice is associated with the peripheral route to persuasion, which means that the evaluation of the advice is mainly based upon peripheral cues and not on a careful examination. The TAM perceived usefulness scores also point in this direction.

6. Conclusions

All in all, the experiments indicate that an LKBS is a persuasive message source and that users tend to evaluate the LKBS advice through the peripheral route. In other words: LKBSs may prove to be very persuasive messengers, even when the message is wrong or questionable. Although these experiments were conducted in a laboratory setting, it is likely that 'real' LKBS users will sometimes evaluate LKBS advices in a similar way.

Now, how can we ensure that we construct LKBSs that are helpful to lawyers, when these lawyers have difficulties with making correct evaluations of the usefulness of an LKBS? The user's *perceived* usefulness of an LKBS does not have to reflect the true usefulness of the system. Every LKBS researcher knows that an LKBS is by no means an infallible oracle. Only in the hands of a competent user, an LKBS can be a powerful tool. A competent user should supervise the LKBS and examine whether or not its advices are applicable in a case at hand. If lawyers uncritically adopt LKBSs, they may use them incorrectly and rely too much on the advices given by the system.

The solution to this problem may lie in the construction of small, transparent, knowledge-based systems designed to help lawyers with routine decision making. Improving the user-friendliness of the LKBS is not enough when users are not motivated to supervise the system and rather use them to reduce their own cognitive effort. LKBSs should be conceptualised as instruments, tools that supports effective performance. And sometimes an LKBS is just not the right tool, like for instance Lodder, Oskamp and Duker [25] concluded in a project for the Dutch Victim Support Organisation.

The results of the experiments also underline that if we want to put LKBSs into legal practice, we need to study how lawyers will respond to them. Using the ELM and TAM to examine the interaction between users and LKBSs seems to be a good start. However, current knowledge of user behaviour is limited and more research is needed to fully understand how users can be influenced by computerised decision support in general and by LKBSs in particular.

References

[1] Dijkstra, J.J., On the use of computerised decision aids: An investigation into the expert system as persuasive communicator. Unpublished dissertation, University of Groningen, Groningen, 1998.

[2] Aldag, R.J., and D.J. Power, An empirical assessment of computer-assisted decision analysis, *Decision Sciences* **17** (1986) 572-588.

[3] Davis, F.D., *et al.*, Harmful effects of seemingly helpful information on forecasts of stock earnings. *Journal of Economical Psychology* **15** (1994) 253-267.

[4] Davis, F.D., and J.E. Kottemann, User perceptions of decision support effectiveness: Two production planning experiments, *Decision Sciences* **25** (1994) 57-77.

[5] Kottemann, J.E., *et al.*, Computer-assisted decision making: Performance, beliefs, and the illusion of control, *Organizational Behavior and Human Decision Processes* **57** (1994) 26-37.

[6] Nielsen, J., and J. Levy, Measuring usability: Preference vs. performance, *Communications of the ACM* **37** (1994) 66-75.

[7] Petty, R.E., and J.T. Cacioppo, Communication and persuasion, central and peripheral routes to attitude change. Springer Verlag, New York, 1986.

[8] Petty, R.E., and J.T. Cacioppo, The elaboration likelihood model of persuasion. In L. Berkowitz (ed.), Advances in experimental social psychology. Academic Press, New York, 1986, pp. 123-205.
[9] Eagly, A.H., and S. Chaiken, The psychology of attitudes. HBJ, Fort Worth, 1993.
[10] O'Keefe, D.J., Persuasion: Theory and research. Sage Publications, Newbury Park, 1990.
[11] Davis, F.D., *et al.*, User acceptance of computer technology: A comparison of two theoretical models, *Management Science* **35** (1989) 982-1003.
[12] Davis, F.D., Perceived usefulness, perceived ease of use, and user acceptance of information technology, *Management Information Systems Quarterly* **13** (1989) 319-340.
[13] Payne, J., *et al.*, The adaptive decision maker. Cambridge University Press, Cambridge, 1993.
[14] Todd, P. and I. Benbasat, The use of information in decision making: An experimental investigation of the impact of computer-based decision aids, *Management Information Systems Quarterly* **16** (1992) 373-393.
[15] Todd, P., and I. Benbasat, The influence of decision aids on choice strategies: An experimental analysis of the role of cognitive effort, *Organizational Behavior and Human Decision Processes* **60** (1994) 36-74.
[16] Todd, P., and I. Benbasat, The influence of decision aids on choice strategies under conditions of high cognitive load, *IEEE Transactions on Systems, Man, and Cybernetics* **24** (1994) 537-547.
[17] Duan, Y., *et al.*, An analysis of users' reactions to an expert advisory system. *Expert Systems with Applications* **9** (1995) 271-282.
[18] Hill, T., *et al.*, Role of efficacy expectations in predicting the decision to use advanced technologies: The case of computers, *Journal of Applied Psychology* **72** (1987) 307-313.
[19] Dijkstra, J.J., The influence of an expert system on the user's view: How to fool a lawyer, *New Review of Applied Expert Systems* **1** (1995) 123-138.
[20] Dijkstra, J.J., *et al.*, Persuasiveness of expert systems, *Behaviour & Information Technology* **17** (1998) 155-163.
[21] Kiers, H.A.L., SCA user's manual. ProGAMMA, Groningen, 1990.
[22] Murphy, D.S., and S.A. Yetmar, Auditor evidence evaluation: Expert systems as credible sources, *Behaviour & Information Technology* **15** (1996) 14-23.
[23] Dijkstra, J.J., User agreement with incorrect expert system advice, *Behaviour & Information Technology* **18** (1999) 399-411.
[24] Hart, S.G., and L.E. Straveland, Development of NASA-TLX. In P. A. Hancock and N. Meshkati (eds.), Human mental workload. North-Holland, Amsterdam, 1988, pp. 139-183.
[25] Lodder, A.R., *et al.*, AI & criminal law. In: H.J. van den Herik et al. (eds.), Legal knowledge based systems – JURIX 1998: the eleventh conference. GNI, Nijmegen, 1998, pp. 59-70.

Appendix A

An example of a criminal attempt case used in the first and the second experiment (from [19]). The original text was in Dutch.

Motorcycle case

Mr A has been annoyed for weeks by the noise of the motorcycle of his neighbour. Therefore he steals the motorcycle, when he sees it abandoned in the neighbour's backyard. What he did not know was that his wife had bought him this motorcycle a few hours ago.

Should A be punished for the attempt of theft?

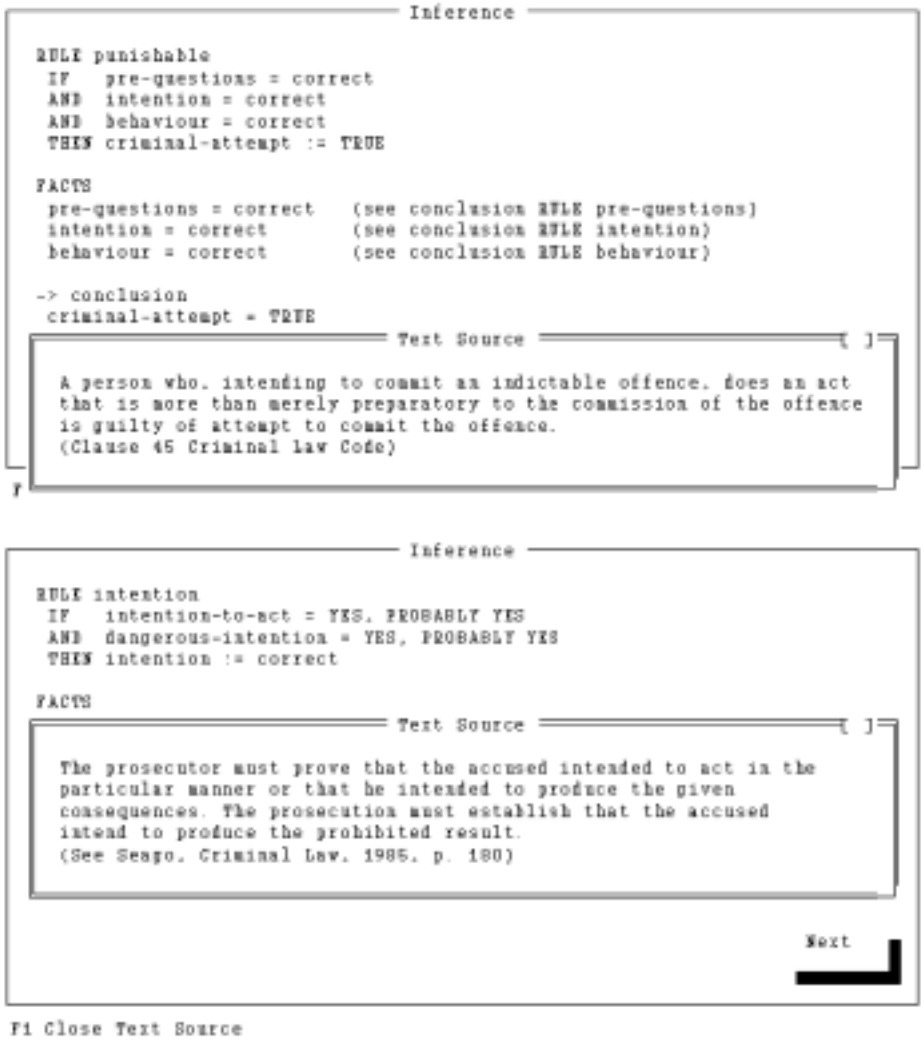
Alternatives:
a) Yes, A should be punished for the attempt of theft. Though he could not steal this motorcycle, because he owned it already, motorcycles are normally suitable objects to steal.

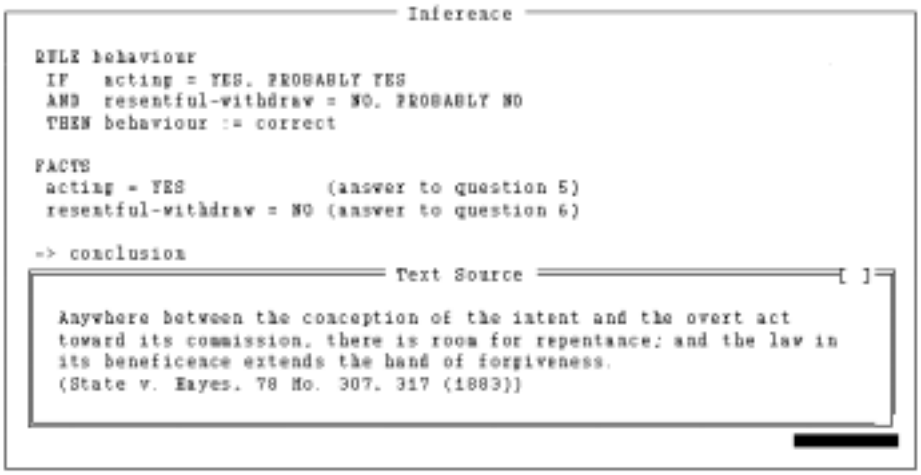
b) Yes, A should be punished for the attempt of theft. He had the intention to steal the motorcycle, and he acted that way. The fact that he already was the owner, made it impossible to commit theft, but it does not make the attempt uncriminal.

c) No, A should not be punished for the attempt of theft. It is impossible to steal something that you own. Therefore it is also impossible to make an attempt to steal something that you own.

Appendix B

Three screenshots from the subjectivistic LKBS. All screenshots are from the how-function with the text-function activated (from [19]). The inferences are from the motorcycle case (see Appendix A).





F1 Close Text Source