World of Uncertainty

Track record

This proposal is being submitted by 3 researchers who came together for the first time at the EPSRC Ideas Factory sandpit on Scientific Uncertainty and Decision-making. Since this is a new idea, generated at the sandpit, we have not worked on it before: in fact, no-one has. However, we have researched areas relevant to this proposal. We have complementary skills, and work in department with complementary strengths. Between us we have the skills needed to run a research project that creates ways of explicitly modelling uncertainty, incorporates them into a computer game, and evaluates the effects of the game upon learners.

Queen's University Belfast

Dr. David Newman, of the School of Management and Economics at Queen's University Belfast, has an MA in Natural Sciences, a Ph.D. in Appropriate Technology for the Third World and an M.Sc. in Artificial Intelligence. The M.Sc. dissertation, on AI techniques to model the expert assessment of uncertain data for biomass energy planning, drew on his experience as a researcher in Mauritius, Mozambique and Kenya.

At Queen's University he has researched groupware use in co-operative learning, critical thinking in on-line and face-to-face discussions, and the use of the Internet by community groups. He developed techniques to measure the depth of critical thinking in collaborative learning, and used them to compare face-to-face and computer supported seminars (Newman 1996). He then changed his teaching techniques to take account of this research, and won a university teaching award for the integration of educational research and pedagogic practice. Later he ran the evaluation component of an EU eContent project, Add-Wijzer (www.addwijzer.info) in which he evaluated the usability and usefulness of on-line legal databases for non-lawyers, in collaboration with public and private sector partners in three countries. He is currently running a €0.5 million cross-border research project into electronic public consultation, funded under the EU Peace II programme for Peace and Reconciliation in Ireland (www.e-consultation.org).

His works within the Information Systems Research Group in the School of Management, which has strengths in IS evaluation. One of his colleagues used narrative theory to study computer game playing for her Ph.D.

John Morison and David R. Newman (2001) "On-line Citizenship: Consultation and Participation in New Labour's Britain and Beyond". *Int. Rev. of Law, Computers & Technology,* **15**(2), 171-194.

Newman, D. R., & Emerson, P. J. (1997). "The on-line preferendum: a tool for voting, conflict resolution and decision-making." In R. Alton-Scheidl, R. Schmutzer, P. P. Sint, & G. Tschertau (Eds.), *Voting, Rating, Annotation–Web4Groups and other projects: approaches and first experiences* (OCG-Bd.104, pp. 131-46). Vienna-Munich: R. Oldenbourg and Austrian Academy of Sciences.

Newman, D. R., Chris Johnson, Brian Webb and Clive Cochrane (June 1997) 'Evaluating the Quality of Learning in CSCL'. *J. Am. Soc. for Information Science* **48**,(6), 484-95.

Newman, D. R. (1989) *A Consultant's Assistant: knowledge-based assessment of uncertain data for biomass energy planning.* M.Sc. in Artificial Intelligence, School of Information Systems, Kingston Polytechnic.

University College London

Professor Philip Dawid has spent his career in the Department of Statistical Science, University College London. He has a broad and deep understanding of the Bayesian approach to statistical modelling, inference and decision theory, and of its potential to contribute to the quantification and communication of personal and scientific uncertainty. His joint book, *Probabilistic Networks and Expert Systems,* won the 2002 DeGroot Prize for a Published Book in Statistical Science. He has long-standing research interests in the construction and use of proper scoring rules and related formal devices to encourage and reward honest assessment of personal probabilities.

He has had a special interest in the communication, uses and misuses of probability and statistics in the law, and has made significant contributions, both conceptual and technical, to the evaluation of forensic DNA identification evidence.

He constructed and directs a cross-Faculty interdisciplinary research programme on 'Evidence, Inference and Enquiry' at UCL, involving (inter alia) Statistics, Economics, Law, Crime Science, Psychology, History, Medicine, Science and Technology Studies, and Education, with ~£1M Leverhulme/ESRC funding.

Dawid, A. P. (1986). Probability Forecasting. *Encyclopedia of Statistical Sciences* vol. 7, edited by S. Kotz, N. L. Johnson and C. B. Read. Wiley-Interscience, 210–218.

Dawid, A. P., DeGroot, M. H. and Mortera, J. (1995). Coherent combination of experts' opinions (with Discussion). *TEST* **4**, 263–313.

Dawid, A. P. (2002). Bayes's theorem and weighing evidence by juries. In *Bayes's Theorem*, edited by Richard Swinburne. *Proc. Brit. Acad.* **113**, 71–90.

Brunel University

Dr Melissa Cole is currently employed as a lecturer in information systems at Brunel University. Information Systems & Computing is one of the largest centres for research in its field in Europe and achieved a grade 5 for research in the last RAE in 2001. It is home to around 70 research active members of staff including 8 professors covering a wide range of research areas. These areas are organised in 3 research themes: Information Systems, People & Interactivity, and Software Technologies & Modelling.

In May 2005, she was awarded a Ph.D.: "A Hermeneutic Investigation of Online Consumer Decision-Making". She has extensive experience of interface design and their role in supporting user decision-making. Her current interests involve the design of interactive software tools that promote reflective practitioners.

R.M. O'Keefe, M. Cole, P.Y.K. Chau, A. Massey, M. Montoya-Weiss and M. Perry. (2000) "From The User Interface To The Consumer Interface: Results From A Global Experiment", *International Journal of Human-Computer Studies*, (53), 611-628.

J.L.Thompson and M. Cole., (1997) "Strategic Competency - The Learning Challenge", *The Journal of Workplace Learning*, 9(1), 45-62

Proposed research

The idea for this research arose during the Ideas Factory sandpit on Scientific Uncertainty and Decision-Making, held at the Shrigley Hall Hotel in the last week of January 2006. These sandpits are events designed to produce research ideas that would be considered too risky to be acceptable when presented to a panel in the normal way.

Our idea is to build and test an educational computer game in which decision-makers can learn how to handle uncertainty.

By experiencing a world of uncertainty, and by explicitly estimating uncertainty in that world, they can develop the skills and attitudes needed to make decisions under uncertain conditions, instead of trying to eliminate uncertainty from the problem space.

The problem context

Consider the three quotations, on the right, from that famous philosophical treatise, *The Science of the Discworld* (Pratchett, Stewart & Cohen 1999). In Fig. 1, Ponder Stibbons, a researcher, is trying to persuade the Arch-Chancellor and senior faculty of Unseen University that creating a world without magic (our world, the roundworld) is a safe experiment. Like many decision-makers in our world, the wizards misunderstand the estimates of uncertainty, applying personal analogies rather than formal risk assessment.

Faced with such misconceptions, in Fig. 2, Ponder, like many scientists, gives up trying to make explicit his model with all its uncertainties, and creates little stories for the decision-makers. In our evolution, humanity have had told and listened to stories for tens or hundreds of thousands of years. Narrative forms are the currency of our entertainment (in films, television, books and magazines) and the way we often think about problems and their solutions. Politicians, journalists, public relations people, barristers, juries and most citizens relate to simple stories, to compelling narratives. Much of the public debate about science is a dramatised argument between two competing narratives. In these discourses, be it evolution versus intelligent design, or global warming versus economic growth, the subtleties and uncertainties are squeezed out in creating simple, understandable, stories. This is so commonplace it has become the subject of parody, as in a sketch broadcast on 4 March 2006 in Bremner, Bird and Fortune (see Fig. 3).

Now that is how the Discworld works. Stories drive events, people act in ways to confirm the story: so much so that it takes all the efforts of three powerful witches to stop Emberella marrying the prince (Pratchett 1992). On the Discworld, million-to-one chances crop up nine times out of ten (Fig. 4). But that is not how our world works. Ours is a world of uncertainty, where truth is stranger than fiction, and where the common sense story may be false. Just because one person's story appears in the Daily Mail, it doesn't mean that everyone else shares the same experience, or wants the same decision. One woman's miracle drug is another woman's poison. Decisions can be distorted by our natural instinct to believe in stories. That is why PR and spin can be so successful. It isn't a case of lying with statistics, but the narrative lie is better understood than the statistics.

So how can we get more people involved in decision-making to understand uncertainty? Instead of trying to eliminate uncertainty, how can they explore uncertainty, handling it with the aplomb of a bookmaker or professional gambler? If we can get decision-makers and the public to better understand uncertainty, and scientists reporting to them to better understand stories, then we stand a chance of bridging this knowledge transfer gap.

'Well ... in the unlikely event of it going seriously wrong, it ... wouldn't *just* blow up the university, sir.'

'What would it blow up, pray?'

'Er ... everything sir.'

'Everything there is, you mean?'

'Within a radius of about fifty thousand miles out into space, sir, yes. According to HEX it would happen instantaneously. We wouldn't even know about it.'

'And the odds of this are ... ?'

'About fifty to one, sir.'

The wizards relaxed. 'That's pretty safe. I wouldn't bet on a horse at those odds,' said the Senior Wrangler. There was half an inch of ice on the *inside* of his bedroom windows. Things like this give you a very personal view of risk.

Fig.1. Decision-makers' understanding uncertainty (Pratchett et al. 1999, p. 20)

Ponder had invented a little system he'd called, in the privacy of his head, Lies-to-Wizards. There was no *point* in telling your bosses *everything*; they were busy men, they didn't want *explanations*. There was no *point* in burdening them. What they wanted was little stories that they felt they could understand, and then they'd go away and stop worrying.

Fig. 2. Scientists' communication to decision-makers (Pratchett et al. 1999, p. 16)

'It is impossible for human beings to catch bird 'flu from birds. ... Unless, it is possible for them to catch bird 'flu from birds, in which case, if it's possible for humans to catch bird 'flu from birds, millions of people will die.'

'Yes. And that's the latest scientific advice we have?'

'Yes. Before there was just confusion, and now we have this two clear alternatives ... either it will be fine, or it will be a catastrophe.'

Fig. 3. How decision-makers understand scientific advice (Bremner et al., 4 March 2006)

On the Discworld, it is clearly recognised that million-to-one chances happen nine times out of ten. The reason is that every Discworld character lives out a story, and the demands of the story determine how their lives unfold. If a million-to-one chance is required to keep that story on track, then that's what will happen, appalling odds notwithstanding.

Fig.4. When narrative trumps uncertainty (Pratchett et al. 1999, p. 246)

The solution space: educational games

Where can we explore this world of uncertainty, and develop our skills in estimating and managing it? There is one space where linear narrative does not dominate: the computer game. Unlike films, books, or even TV science documentaries, a computer game does not follow a single narrative thread. Instead the player creates his or her own narratives as a result of his or her decisions and actions while playing. A computer game is necessarily interactive but it is not in itself a narrative. Mallon and Webb (2005) note that theorists have been engaged for some years in the debate over the utility of narrative versus interaction for games, including Laramee, Onder, Juul, Frasca, Jenkins, Ryan, Eskelinen, Costikyan , Egenfeldt-Nielsen, Adams, and Talin.[1] Players do pay attention to narrative elements, such as stories in the world with which they are interacting. But the play is not limited to following a pre-written story.

Inside the environment of a computer game, a player may explore, gather evidence, estimate risks, make decisions, and see the consequences of these decisions. It seems to be an ideal environment for exploring uncertainty. Indeed, Mallon (2004) found that, ‘Players agreed that chance reduces the sense that a story is pre-set. They discussed two types of chance element: (a) arbitrary randomness and (b) random elements resulting from the spontaneous combination of a range of variables, within which players can strategically pre-plan and manage their resources. The introduction of chance elements from within the second category was most favoured and elements from the first category were contentious.’

Lest the use of games as a resource for discovering evaluation or design criteria seem frivolous, it is worth remembering Bielenberg & Carpenter-Smith’s (1997) claims as to the common goals of entertainment and education, or Ju and Wagner’s (1997:79) point that a computer game is educational because it “trains problem solving skills and fosters learning” and reasoning, including decision making or resource allocation. Malone (1981) investigated games to develop mechanisms for enhancing intrinsic motivation in instructional design. Furthermore, Murray (1997: 144) asserts that game playing skills “have always been adaptive behaviors”. While classified as recreational “because they offer no immediate benefit to our survival … Games traditionally offer safe practice in areas that do have practical value; they are rehearsals for life.”

One difference between an educational game, and a game designed for entertainment is that an educational game is more explicit. The learners do not only internalise skills through practice: they also receive feedback on the content of their learning and their performance. After playing the game they should have both a kinaesthetic and cognitive appreciation of the skills and knowledge they have learned. We need to find a way of designing a game that is still playable, but makes uncertainty explicit. Players need a sense of control over the game, so uncertainty should be something they can affect by actions, or at least estimate and allow for, not just random ‘shit happens’. Indeed, there is a hypothesis that challenge depends on goals with uncertain outcomes (Malone, 1981; Kagan 1978:157; Eifferman, 1974). As players develop their skills, they take on bigger challenges. Optimal conditions for engagement (or ‘flow’) occur when there is: ‘a sense that one's skills are adequate to cope with the challenges at hand, in a goal-directed, rule-bound action system that provides clear clues as to how well one is performing.’ (Csikszentmihalyi 1990) So how can we create a game which challenges players to improve their skills of handling uncertainty?

Encouraging careful uncertainty assessment

The core function of an uncertainty game is to encourage players to deliberate carefully about their true uncertainty and to attempt to quantify it honestly and accurately. In order to do this, players must be rewarded according to how well they conduct these tasks.

The scenarios we have in mind are closer to horse racing than card games. Different punters will have different views as to how likely it is that Dark Lady will win the 2:30 PM race at Doncaster, and different bookies may offer different odds on this event. There is no "true probability", and even in the same state of information different individuals will have different uncertainties. We cannot reward an individual for being "right" or "wrong" about the value of this probability. Nevertheless, we do have an intuition that some values are more reasonable than others (again, in the light of specified information). And certainly some bookies are better than others.

When we come to assessing the value of a probability forecast, or forecaster, we need to distinguish two different dimensions: substantive and normative skill. Substantive skill refers to expertise in the area, e.g. horse racing or weather forecasting. However, if the meteorologist has a good feel for the uncertainty in the weather but is bad at putting it into usable numbers, he has poor normative skill.

Suppose that a player is asked to input her value for the probability of some event. When there is a "right answer" for the value of a probability, the basic task is obviously to compare the true value with the input value, to see how close to player got. However, we cannot do this when there is no right answer.

In this case, we must somehow try and compare the input value (which is a number between 0 and 1) with the actual outcome of the event (which can be coded "true" or "false"). And as these appear to inhabit different universes, such a

1 See the Mallon and Webb articles for the full arguments and references on games, narrative and interaction, as well as empirical data from Mallon’s Ph.D. thesis on how players evaluate and interpret computer games.

comparison would appear problematic. However, progress can be made. The fundamental idea is to construct a scheme of rewards for the player, where the reward or penalty obtained depends on both the input probability and the eventual outcome of the event whose probability is being assessed. A good scheme of this nature would have the following properties:

1. It encourages honesty -- when the player truly believes that the probability is P, her best choice of probability to quote, Q, is Q=P.
2. Over a reasonable sample of occasions, of two players with the same expertise (substantive skill), the player with the better normative skill will earn the higher reward.
3. Likewise, of two players with the same normative skill, the player with the greater expertise will earn the higher reward.

There are in fact a number of schemes that have these properties. General theoretical background may be found in Dawid (1986). There are several possible ‘proper scoring rules’. An example is the rule that provides for a penalty of $(1 – Q)^2$ if Dark Lady wins, as against Q^2 if she does not. If your true probability is P, then your best choice for Q, in the sense of minimising your expected loss, is P – again, encouraging honesty. And once again the accumulated score, over a number of similar occasions, will have properties 2 and 3 above. This particular proper scoring rule is only one of a very large family. An appropriate rule can be constructed to focus on the important aspects of any real decision problem that may be faced.

We can also decompose a player's cumulative score, over many occasions, into two component parts, one measuring substantive skill, and the other normative skill. A player is said to be "well-calibrated", if, over a series of events, close to 70% of those events to which she attached a probability forecast of 70% in fact occurred (and similarly for other probability values). It is valuable to perform such comparisons, which measure normative skill. For a poorly calibrated forecaster, this may assist us and her to attach more realistic numerical values to probabilities in the future. Immediate feedback is the ideal. A player should be able to assess her probability for one event at the time, then observe the outcome of the event, and reap the appropriate reward or penalty before going on to make a probability forecast for the next event.

This scoring and feedback would be part of an uncertainty model (or uncertainty engine) that could be incorporated into many different kinds of game: from a simple quiz game (in which players gave not only their answers, but also how uncertain they felt), through sequential problem-solving simulations, right up to MMORPGs (Massive Multiplayer On-line Role-Playing Games).

Educational scenarios and cognitive outcomes

The above considerations take no account of the psychological biases that human players are subject to. These have been intensively studied (Kahneman et al., 1982), and include such effects as the availability, anchoring, and representativeness biases. Psychological findings can also assist in presenting information in ways that minimise the distorting effects of such biases.

If a computer game is going to have an effect on learners, it will have train players to overcome such biases. These are particular critical thinking skills, which players will need to use in combination with the common problem-solving skills laid out in the educational literature, such as in Garrison’s Theory of Critical Thinking (Garrison 1992). Newman et al (1997) have developed ways of measuring critical thinking in learning situations. We can extend these evaluation techniques to measure critical thinking about uncertain situations.

To properly evaluate educational and psychological effects on learners, we will need to test the game in a realistic educational setting. For example, suppose we develop a game in which players explored evidence on threats of a bird ‘flu epidemic, estimated risks, made decisions (in teams) and then saw the consequences. Then the ideal test setting is with veterinary students, learning to handle epidemics.

The intention is to create the game as a framework for structuring a particular type of scenario, for example, the strategic management of an unfolding crisis situation where a virus, pollutant or flooding is spreading through a geographical area. The “player”, who only sees this through information reaching a control point, would “win” not by simply stopping the disaster but by being able to comprehend and manage the uncertainties in the information to hand and the outcome of their actions. This framework would then have the potential to represent different scenarios, appropriate to the trainee’s area of interest, by providing different controlling data sets.

The game design needs to specifically consider how it motivates the player to participate and become immersed in the scenario. In the leisure market graphic realism and sound effects are believed to have a significant impact. However, in this case the “player” is, in part, motivated to participate by the desire or need to improve their ability to handle uncertainty. Realism, or believability of the scenario and the way it is perceived, is thus likely to be more critical than surface visual presentation. Although attractive graphical elements are important full 3D rendering of moving environments are of less significance than the information content.

The project

Commercial games can cost $10 million dollars to develop. So what can we do within the limitations of one small research project? We cannot develop a commercial game for Playstation 3 or Xbox 360, but we **can** design and develop a small prototype game as a proof of concept, and evaluate its educational effects. Our objectives are to:

1. Design an explicit model of uncertainty that could be included in any computer game.
2. Investigate the educational, psychological and mathematical requirements for learning about uncertainty.
3. Identify a range of game scenarios in which decision-makers might learn about uncertainty.
4. Develop and test a simple game for one of these scenarios.
5. Evaluating the effects of this game on a sample of learners.

Methodology

This is a research project which incorporates within it a software development process. The research and development methodologies are closely linked. The approach is that of Human-Centred Systems. The systems design, the development process, and the evaluation are shaped by human needs: the needs of its intended users, potential decision-makers learning about uncertainty. The Human-Centred Systems approach includes the workplace enhancement activities of the Tavistock Institute, as embodied in Enid Mumford's ETHICS methodology, Soft Systems Methodology, and modern user-centred approaches to interaction design. In brief, we will start from the needs of the learners, and use these needs to shape the design of the game, its initial evaluation, and the evaluation of its effects on learners.

The game will be developed using evolutionary delivery, which is an iterative approach to software development that negotiates and delivers increased functionality in each cycle. This methodology is appropriate where requirements are not completely pre-specifiable enabling the game to be explored and evaluated throughout the project in a continuous manner without having to wait for classic development stages to complete. The developer alternates between “the requirements phase” and the “design phase” as desired. What this means for the research project is that essential but low-effort story-boarding activities can occur around effort-intensive activities focused on creating game parameters and/or determining core functionalities. It is anticipated that such activities will form one round of iteration involving 2-3 months of intensive effort. Such iterations will themselves, be driven by the findings and outputs related to scenario development and the Bayesian probability engine. It is envisaged that the scenarios will be multi layered and the game will include different levels of complexity. This will be achieved through evolutionary developments that incorporate evaluations of user perceptions of risk and uncertainty and an increasingly sophisticated Bayesian probability engine.

This development methodology sits inside an overall user-centred evaluation methodology. A successful project would produce a game that is usable by the intended learners, is useful in developing their skills in and knowledge of uncertainty, and has measurable qualitative (narrative) and quantitative (cognitive) effects on their learning.

Programme of work[2]

Activities 1: understanding the users

Before starting to build any game we need to understand two things: (a) what decision-makers, learners and their teachers would require in order to effectively improve understanding of uncertainty; and (b) possible scenarios of use, identifying decision-making contexts for which a game might be developed, and groups of people who might use such a game to learn about uncertainty.

This would be carried out by a Ph.D. student, supervised by Dr. Newman. It will start with a literature review: including the literatures on behavioural decision-making, learning styles and critical thinking, educational games, narrative and games, and studies focussed on learning about uncertainty. (a) will require a survey of decision-makers to find out what problems they had in coping with uncertainty, and telephone interviews with educators to explore strategies for developing skills and understanding of uncertainty. From the same sources, we would identify a series of scenarios, possible game worlds and pedagogic applications.

Activities 2: the uncertainty engine

Since this is an educational game, we need to find ways of explicitly modelling and representing uncertainty that could be implemented in an uncertainty engine within a computer game. This uncertainty engine would receive input from the game (e.g. users entering their uncertainty), and return responses for feedback to the players (e.g. scores, performance charts, changing parameters that affect the next problem to be solved, and full explanations). It would go beyond the current Bayesian models used in some games, as it would make its workings transparent and explicit: it is for deliberate interaction, not random variation. It could, in principle, be incorporated in many different games, used like physics models or AI models are in the current generation of computer games.

2 Note that the resource justification is in a separate annex, as requested by EPSRC.

This work would be started by the Ph.D. student, working to develop Prof. Dawid's ideas for Bayesian scoring systems into algorithms and heuristics, and then, working with the games developer, turn that into an executable uncertainty engine. The exact contents of the uncertainty engine will be defined at this stage. We anticipate that it would at least support calculations of probabilities for the game world, collect and maintain data on players' subjective estimates of uncertainty, and give explicit and implicit feedback on their performance (as scores and reports).

Activities 3: the prototype game

As a proof of concept, we need to build and test a computer game. We will pick one of the game scenarios identified in Activity 1. It will be chosen as a compromise between the ease of development of a game (since we have limited resources) and the potential for large-scale use and dissemination of the prototype game. It must also have a defined pedagogical context for evaluation of the effects of the game on learners, and make use of the uncertainty engine (so that this is an explicit treatment of uncertainty).

An experienced games developer will storyboard and develop the game, using an evolutionary delivery approach. He will be work as a contractor to Brunel University, supervised by Dr. Cole, in collaboration with the Ph.D. student and teachers and learners interested in the use scenario. As each part of the game is developed, it will be tested (for functionality, usability and playability).

Activities 4: evaluation

There are several different types of evaluation activities in this project. There is functional/system performance evaluation, usability evaluation, evaluating the playability and the player-constructed narratives, and evaluating the learning effects (on critical thinking, skills at exploring and estimating uncertainty, and understanding of and attitudes to uncertainty).

From the beginning, in activity 1, the Ph.D. student will research evaluation methodologies consistent with the elicited user requirements. Once the game scenario is chosen, these will be refined into more specific evaluation techniques.

During the game development, each incremental prototype will be tested for functionality, then tested for usability. Once a fairly complete game is ready, we will run playability tests, evaluated both quantitatively (factors engaging or disengaging the players) and qualitatively (from game narratives generated by focus groups of players). The prototype will be modified to make sure the game is playable, usable and functional before educational evaluation starts.

The final evaluation stage is to study the effects of playing the game on learners. The game will be used in a suitable pedagogic context (e.g. veterinary students learning about decision-making in epidemics, or participants in decision-making courses at the National School of Government, or the Queen's University Institute of Governance). The evaluation will be designed in the context of learner requirements, collecting evidence both from the immediate actions during game play (on the psychology of uncertainty estimation), and from learner reflections after having completed a module or course in which they can compare the effects on their learning of the game and other learning experiences (on deep learning, critical thinking and problem-solving processes).

Beneficiaries, dissemination and exploitation

Ideas Factory sandpits are designed to generate new, risky, ideas. There is consequently no guarantee that people will benefit or that something will be produced that can be disseminated or exploited. However, since they are innovative ideas, if everything goes well, then the potential benefits can be enormous. EPSRC is betting on a long shot.

Our vision is of a society transformed from one in which most people prefer simple stories, and avoid discussing uncertainty, to one where a large proportion of the population has the skills of exploring uncertain evidence and can estimate uncertainty. This project will not, on its own, transform society. But it may show a way to improve the education and training of current and future decision-makers, so that they are better able to handle uncertainty in their decision-making. It could reduce the knowledge transfer gap between those happy to think in terms of quantitative models, risks, uncertainty and statistics, and those happiest working with logic, words, and narrative. At present the subtle and deep knowledge of the scientist is over-simplified to make it comprehensible to politicians, journalists, lawyers and non-scientific civil servants.

The veterinary scientist, when presented with an ailing, treasured family pet, or a herd of animals upon which the owner's livelihood depends, faces continual pressure from many directions to produce a confident, definitive diagnosis. A useful educational context in which this game could be used, therefore, is in the training of university veterinary science students; the most effective practising veterinary scientists have a broad understanding of the many and varied uncertainties inherent in their profession. The approaches put forward in this proposal provide a novel and hands-on approach by which concepts of scientific and diagnostic uncertainty are explicitly built into the diagnosis process; the "gaming scenario" is not dissimilar to the making of a diagnosis in a real clinical situation, or in crisis decisions (such as in epidemics), and could in principle be used to great advantage from the very outset of a veterinary training.

Fig. 1. An uncertainty game scenario: vets and epidemics

Some of the meaning is lost, hindering decision-making. In particular, we hope to develop games that help people develop and practice their skills in estimating uncertainty, while enjoying the learning experience. The very first beneficiary of our project will be the group that agrees to work with us in a trial of the game, in a particular learning scenario (e.g. meteorologists or vets, see Fig. 5). From the early of the project we will establish a user group of interested stakeholders, including both those who will take part in the evaluations, and other who are interested in exploiting the approach after the project finishes in their own learning scenarios.

Once we have proved the concept, what can be disseminated? There is a chance that the computer gaming industry will take to explicit uncertainty engines as they have done to physics and AI models. Our initial uncertainty engine could be the start of a trend. It is unlikely that there would be so much interest that the games companies would pay for the engine. In any case, the mathematics of Bayesian statistics are well known. So we instead propose to release the uncertainty engine under a free software licence, such as LGPL. This means that others can adopt and adapt the code quickly, leading to much more rapid development than we can manage with the resources of one research project. We would also present the engine at conferences attended by games developers, to encourage rapid take-up.

The research findings, particularly those on the educational and psychological effects of the game on the players, will be disseminated to researchers in the usual ways, through academic journals and conference presentations. To reach practitioners, particularly decision-makers, we will set up a mailing list from the beginning of the project, through which they can follow the progress of the project from the beginning, and present our findings at workshop at the end of the project.

The prototype game itself is not likely to have the graphic quality expected of commercial games. We do not have $10,000 to invest in that. However, it serves as a proof of concept, a model for future game development, funded by stakeholders in scientific decision-making, or games companies. There is a slight possibility of a spin-off company developing storyboards or games prototypes for stakeholders.

References

A number of these referenced works include bibliographies of articles dealing in more detail with the issues discussed in this case for support.

Bielenberg, D.R. and Carpenter-Smith, T. (1997) Efficacy of story in multimedia training. *Journal of Network and Computer Applications*, *20,* 151-159.

Bremner, R., Bird, J., & Fortune, J. (Directors) (2006, March 4). *Bremner, Bird and Fortune*. London: Channel 4. (AVI recording attached to submission, as birdFlu.avi).

Csikszentmihalyi, M. (1990) *Flow: The Psychology of Optimal Experience.* New York: Harper & Row.

Dawid, A. P. (1986) Probability Forecasting. In S. Kotz, N. L. Johnson and C. B. Read (eds.) Encyclopedia of Statistical Sciences, London: Wiley-Interscience. Vol. 7, 210–218.

Garrison, D. R. (1992) Critical thinking and self-directed learning in adult education: an analysis of responsibility and control issues. *Adult Education Quarterly*, 42(3), 136-148.

Ju, E. and Wagner, C., (1997), Personal Computer Adventure Games: Their Structure, Principles, and Applicability for Training, *The DATA BASE for Advances in Information Systems, 28*(2), 78-92.

Kahneman, D., Slovic, P. and Tversky, A., eds. (1982). *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge: Cambridge University Press.

Mallon, B., & Webb, B. (1997) *Evaluating narrative in multimedia*. Paper presented at Design, Specification and Verification of Interactive Systems '97.

Mallon, B., & Webb, B. (2000). Structure, causality, visibility and interaction: propositions for evaluating engagement in narrative multimedia. *International Journal of Human-Computer Studies, 53*, 269–287.

Mallon, B. (2004) Ph.D. thesis. Queen's University Belfast.

Mallon, B., & Webb, B. (2005). Stand up and take your place: narrative elements in narrative adventure and role-play games. *ACM Computers in Entertainment, 3*(1).

Malone, T.W. (1981) Toward a Theory of Intrinsically Motivating Instruction. *Cognitive Science, 4*, 333-369. Murray, J. H. (1997) *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. New York: The Free Press, Simon & Schuster Inc.

Newman, D. R., Chris Johnson, Brian Webb and Clive Cochrane (June 1997) 'Evaluating the Quality of Learning in CSCL'. *J. Am. Soc. for Information Science* **48**,(6), 484-95.

Pratchett, T. (1992) *Witches Abroad.* London: Corgi Books.

Pratchett, T., Stewart, I., & Cohen, J. (1999). *The Science of Discworld*. London: Ebury Press.

Diagrammatic workplan

QUB leads on user understanding and evaluation, UCL on uncertainty, Brunel on game development. The game development tasks are carried out in parallel: storyboard a bit, develop it, test it, then repeat these small steps until we have a playable game ready for educational evaluation over an academic year. The work with the user group continues until the end of the project.