Deliverable Proof – Other document EIT-BP17

KIC project the report results from	Gaming for better decisions under uncertainty
Name of document	Final report
Summary/brief description of document	A description of the operational game prototype summarising the theoretical content of the game, examples and case study.
Date of report	15 December 2017

Supporting documents:

20171213_AMU_EKETraining_ReviewGaming.pdf

CONTENTS

Gaming for better decisions under uncertainty

CLIMATE-KIC IDEATOR PROJECT

FINAL REPORT

Project leader Ullrika Sahlin, Lund University

Project period Oct-Dec 2017

Participants:

Ullrika Sahlin (PI), Lund University, Sweden, Igor Kozine and Melanie Kreye, DTU, Denmark, Anthony O'Hagan, University of Sheffield, UK, Matthias Troffaes, Durham University, UK, Igor Linkov, Carneige Mellon University, US, Boris Schonfeldt, Vejle municipality, Denmark, Hampus Söderström, Den Nya Verkligheten (game developer and programmer), Scott Ferson, University of Liverpool, UK, Rasmus Bååth, Lund University, Sweden, External: Peter Gärdenfors, Lund University, Olaf Mosbach-Schulz, EFSA, Rajiv Vaid Basaiawmoit, Aarhus University, Denmark

Author of this report is Ullrika Sahlin. Any mistake is her.

E-mail: Ullrika.Sahlin@cec.lu.se

SUMMARY

Gaming for better decisions under uncertainty is a project with the goal to develop a game that motivates people to learn about uncertainty analysis and decision making in an entertaining way. The idea was to use a game environment to explain the benefits of expressing uncertainty when making predictions and possible ways to make decisions under uncertainty.

The goal has been achieved. A group of experts have met and identified the content for the game. Three modules cover tasks related to expressing and understanding uncertainty using prediction and decision problems from daily life that players can relate to. The modules has been documented and populated with several short games/apps which is available online.

The fourth module cover fictional game scenario that involves decision making under climate uncertainty, where there is a need to consider the balance between social, economic, and environmental impacts. Multi criteria decision analysis without and with uncertainty is demonstrated on a climate mitigation decision problem on choosing a bus technology. The concepts resilience and robustness to uncertainty is illustrated on a climate adaptation decision problem related to flood management.

The prototypes have been tested on independent test persons. The project has identified several ideas and needs of further development of existing and new game prototypes. A conclusion is that existing initiatives to train experts in expressing and understanding uncertainty can benefit from gamification to gain more interest. A continuation can help to better link to these. There is also a high value in developing the case-studies on climate decision problems under uncertainty further so they can better serve societal decision makers. We conclude that the output from this project can serve as a start for such development closer to decision makers and stakeholders.

The project has been externally evaluated by the Assessment and Methodological Support Unit at EFSA, who conclude that games like this stimulating learning by scientific experts would be useful for organisations working with scientific assessments.

Gamification provide new means of communicating science and it has been both challenging and rewarding to put together games that are entertaining but with a clear goal to introduce a theory or concept. Several needs and ideas for future developments has been identified.

INTRODUCTION

Uncertainty and decisions

Uncertainty is about admitting the limits of what we know. Scientific experts are usually more custom to communicating things that are certain, proven or well-tested. In association to providing decision support, and especially when assessing risks, scientific expert's often need to make judgements (Burgman 2015). These judgments are not only difficult to make, but also subject to cognitive fallacies and heuristics (Tversky and Kahneman 1974). Expert judgement can be seen as a scientific discipline in itself combining psychology and mathematics (probability theory). Today, scientific uncertainty, issues of trust and independence of experts in scientific assessments (like the IPCC) and a call for transparency in the science to policy interface cast light on expert judgement to support decision problems (Pidgeon and Fischhoff 2011).

Decision theory can be descriptive – describing how decisions are being done, normative – telling how decisions ought to be done and prescriptive – telling how a decision should be done in a particular context. Decision theories exists for identifying strategies against a stochastic nature, other decision makers. Decision problems are of different types such as choosing between alternatives, optimising something or finding a safe threshold (Fischhoff and Davis 2014). In addition, decision theories exists for decision making under risk and uncertainty (Gärdenfors and Sahlin 1988).

Decision makers are influenced by many things, where scientific evidence ought to be a key factor. Wherever there is knowledge there is more or less uncertainty which must be considered by both scientific experts and decision makers. We believe that an interest to learn more about uncertainty will result in a better understanding of uncertainty, which will improve how uncertainty is dealt with in societal decision problems and thereby in better decisions. Uncertainty is an object on the boundary between scientific experts and decision makers, offering for example means of communication of what is known and desired and a space for negotiations on what are useful decisions.

There is uncertainty about many science based decision problems. Climate decision problems are roughly divided into climate mitigation and climate adaptation. There is a large need to make decisions under uncertainty in climate decision problems which deals with long term and often irreducible impacts, a lot at stake, conflicting values and scientific uncertainty. We expect the need for adequate training in uncertainty and decision theory as the number of scientific experts working with climate decision problems and the more decision makers adopting strategies targeting climate mitigation and climate adaptation are growing.

Uncertainty in Science

With the goal of societal decisions robust to scientific uncertainty, a start can be to demonstrate the benefits of treating uncertainty in theory and practice and create opportunity for discussions and learning. This means to acknowledge theory of fields such as expert knowledge elicitation, psychology of expert judgement, probability theory, uncertainty quantification, decision theory, uncertainty analysis and uncertainty communication.

In some research programs and education this is will be a long journey to take, especially when it is in conflict with other goals. A clear example is the discussion on what type of statistics to teach in Science, classical statistics, Bayesian statistics or both (Cox 2006). A second is the lack of training to prepare researchers to be part of formal expert knowledge elicitation (O'Hagan et al. 2006), where the researcher need to understand what uncertainty is, how it can be quantified and how to express it well. A third is the role of decision science (Pidgeon and Fischhoff 2011) – if and at what time to introduce decision theory to students in Chemistry? Uncertainty and decision making fits well in an interdisciplinary science context, but all future scientific experts are not necessarily in such environments.

Gamification

This ideation project started with a need to create a channel to reach out to gain more interest in “uncertainty and decision making” to the current and future generations of scientific experts and decision makers. Our solution was to spark an interest with something engaging and entertaining. So why not as a game?

Playing games is fun. However, entertainment is not always the only purpose of a game. For instance, gaming can serve to assist learning, as in e-learning and “serious gaming”. Morford et al. (2014) view gamification as a “process by which non-game activities are represented in a game-like form”. In education programmes it can promote learning motivation and outcomes (Maturo and Setiffi 2016). Yukai Chou with a lot of experience in gamification see gamification as “the craft of deriving all the fun and engaging elements found in games and applying them to real-world or productive activities” (http://yukaichou.com/).

Project aim

This project aims to develop a game that motivates people to learn about uncertainty analysis and decision making in an entertaining way.

Many games naturally exploit uncertainty and decision making under uncertainty as means for entertainment. In this project, these aspects will be made very explicit: players will be introduced to the concept of subjective probability, will be able to practice their skills in decision making under uncertainty, and get explicit feedback about how their decisions relate to theory. The game will evaluate their gameplay on how good they are in making predictions (and expressing uncertainty) on different types of problems. Thereby, the game will explain the benefits of expressing uncertainty when making predictions and possible ways to make decisions under uncertainty.

This game will introduce the terminology and theory of modern uncertainty analysis including relative frequency, subjective probability, interval on probability, and decision rules that are robust under uncertainty. The game will embrace a wide view on uncertainty management, through a wide range of methods for uncertainty analysis, such as those found in the European Food Safety Authority (EFSA) guidance adopted December 2017.

The game will include prediction and decision problems from daily life that players can relate to, and where players can actually test their accuracy and decision performance in hindsight with the help of the game. During the project, we will identify game scenarios that can fit different types of users.

We will also use a fictional game scenario that involves decision making under climate uncertainty. This scenario will aim to explore strategies to adapt to or mitigate climate change under severe uncertainty and climate scenarios. The decision problem will in order to mimic realistic situations be considering the balance between social, economic, and environmental impacts.

METHOD

Target audience

The target audience were set to high school students, requiring no experience yet a cognitive ability to make qualitied judgments. In addition the game can be used at science fairs and teaching in introductory levels. In this way the game can be used in introductory teaching for scientific experts with no or little training or experience in considering uncertainty in scientific assessments and decision makers.

This game is not supposed to be a course, but it can be linked to such. The E-course "Probabilistic Judgements for Expert Elicitation" has been designed for people who will be experts in an expert knowledge elicitation exercise. This is a course that teaches them how to make the kinds of judgements they will be asked for as scientific experts (http://www.tonyohagan.co.uk/shelf/ecourse.html).

Content

The content of the game was identified at two workshops and meetings with internal and external experts.

The fictional game scenario that involves decision making under climate uncertainty was developed based on a climate decision problem at Vejle municipality, Denmark and the project expert on resilience and multi criteria decision making.

The partners drafted the content of the game which was turned over to the programming and game development.

Design

There exists board games for guessing (such as guesstimaster https://www.worldofboardgames.com/guesstimaster) and decision making (such as risk management decisions and urban planning related to flooding e.g. floodville and riskköping developed at Karlstad University Sweden https://www.kau.se/ccs/samverkan-och-motesplatser/risklab).

We decided to develop computer game to be able to reach out to a wider audience and be able to create artificial situations and a nice interface. Instead of one large game we decided to develop several shorter games.

The aim was to be clear what theory or concept being targeted in each game. An ambition was to have a design that optimise for human motivation (http://yukaichou.com/) as opposed to function. As a consequence target theory or concept were not explained inside some of the games.

Development

A skilled programmer and experienced game developer was contracted to turn the drafted content into a game. A testing group consisting of people independent of the project were created to test prototypes of small games.

The programmer focused on developing games about expressing uncertainty and guessing in the software Unity (https://unity3d.com/). Additional prototypes of games which demonstrates a target theory was developed as ShinyR apps (http://shiny.rstudio.com/). ShinyR uses the open source R program originally developed for statistical computing (https://www.r-project.org/).

Format of project output

The output from this project is an operational game prototype together with descriptions of the game, instructions on how to use it, its theoretical content and the climate decision case study.

RESULTS

Overview of the game

The game consists of different modules for which there are several tasks (Figure 1). One of several tasks can be part of a smaller game. The smaller games include or links to texts explaining theory that the game is designed for.

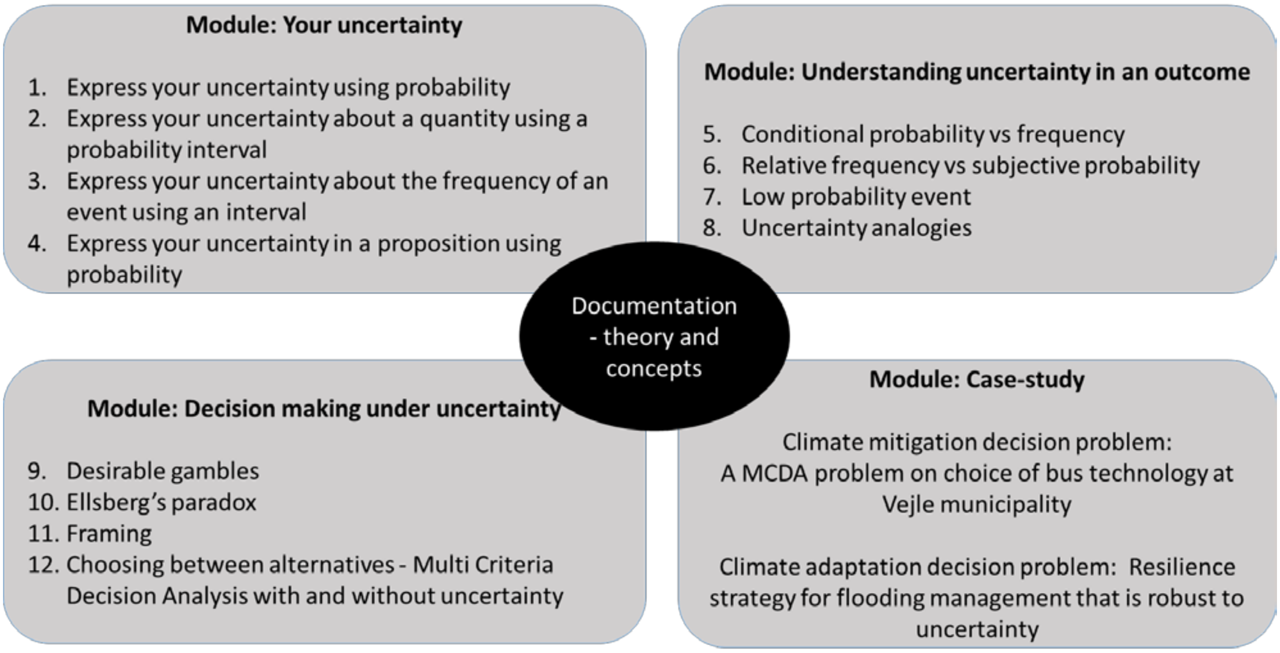
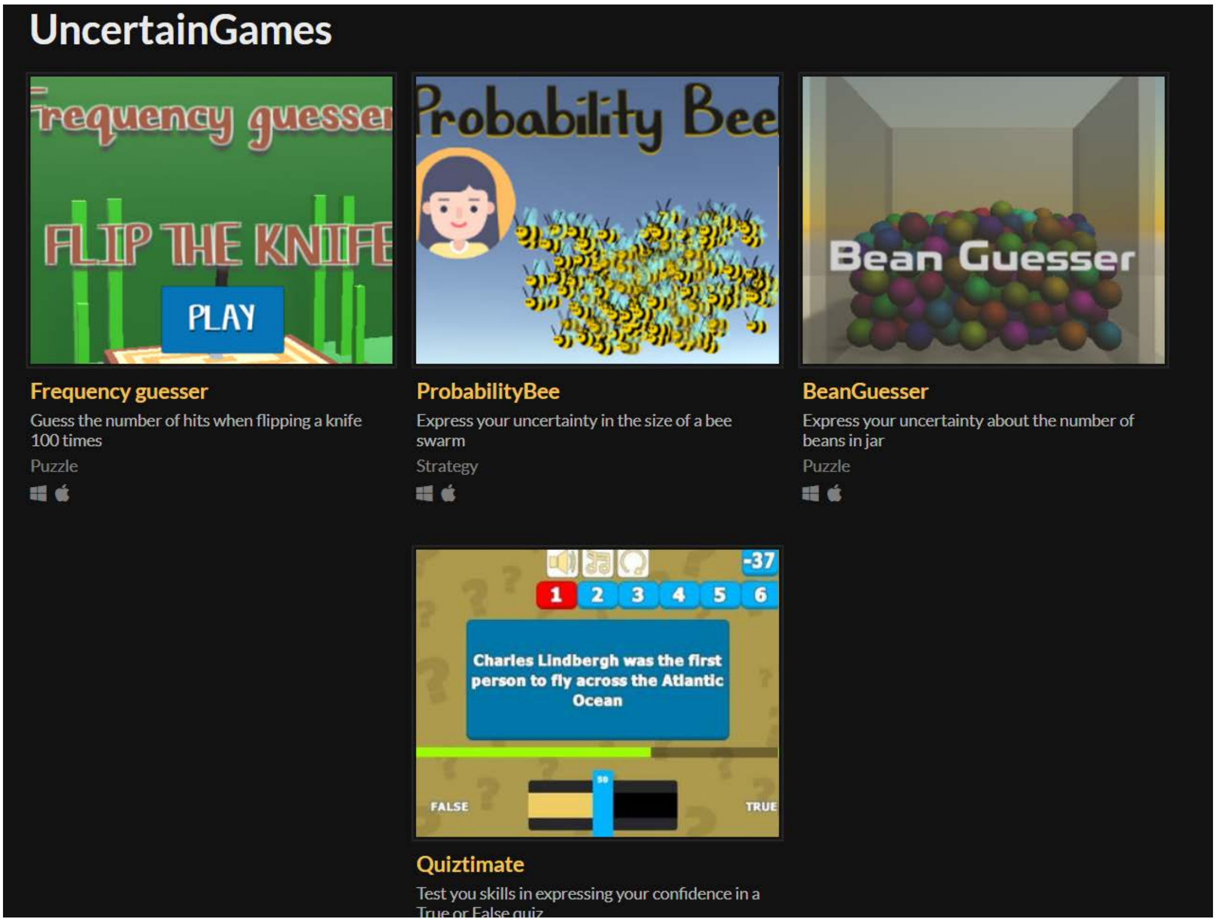
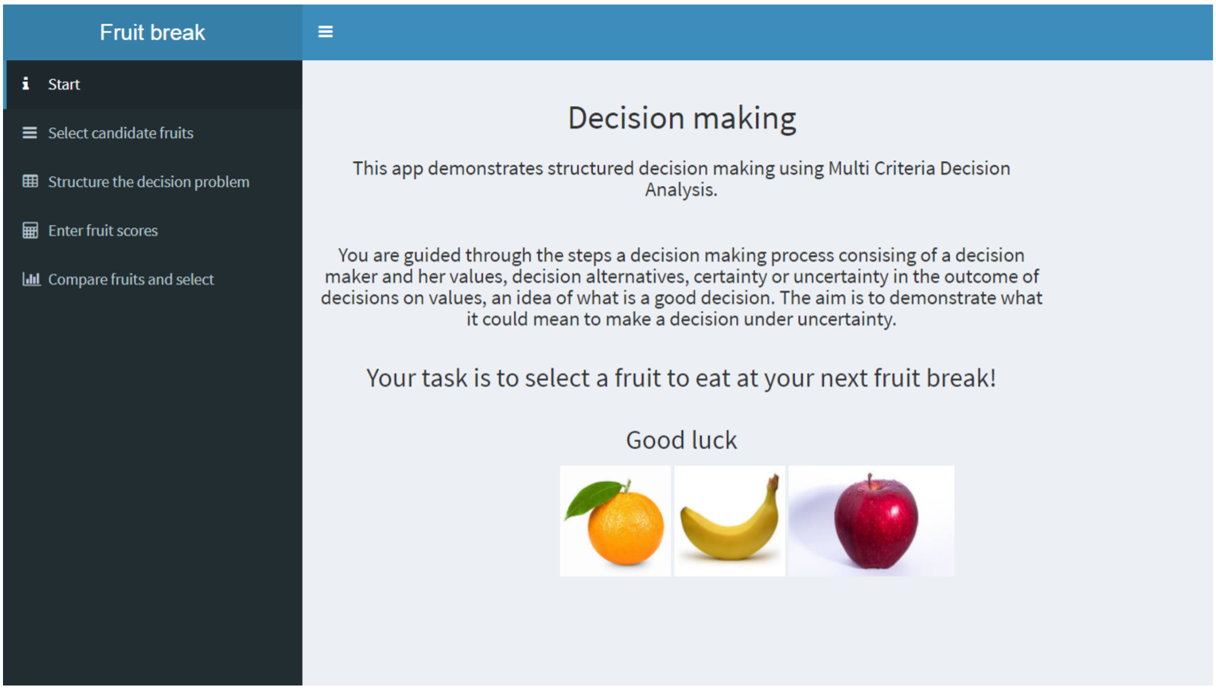


Figure 1. Overview of the game with four modules and tasks.

Daily life problems

Short games, some with high playing appeal and some with a clear aim to demonstrate something, have been developed within the first three modules (Appendix 4). The scoring rules for games targeting expressing uncertainty are described in Appendix 3. Screenshots of games are shown below:


UncertainGames
Frequency guesser
FLIP THE KNIFE
PLAY
Probability Bee
Bean Guesser
Frequency guesser
Guess the number of hits when flipping a knife 100 times
Puzzle
ProbabilityBee
Express your uncertainty in the size of a bee swarm
Strategy
BeanGuesser
Express your uncertainty about the number of beans in jar
Puzzle
-37
1 2 3 4 5 6
Charles Lindbergh was the first person to fly across the Atlantic Ocean
FALSE
TRUE
Quiztimate
Test you skills in expressing your confidence in a True or False quiz


Fruit break
Start
Select candidate fruits
Structure the decision problem
Enter fruit scores
Compare fruits and select
Decision making
This app demonstrates structured decision making using Multi Criteria Decision Analysis.
You are guided through the steps a decision making process consising of a decision maker and her values, decision alternatives, certainty or uncertainty in the outcome of decisions on values, an idea of what is a good decision. The aim is to demonstrate what it could mean to make a decision under uncertainty.
Your task is to select a fruit to eat at your next fruit break!
Good luck

Case studies

Two case-studies have been identified within the fourth module. Each case-study was selected as a fictional game scenario that involves decision making under climate uncertainty, where there is a need to consider the balance between social, economic, and environmental impacts.

1. Multi criteria decision analysis without and with uncertainty is demonstrated on a climate mitigation decision problem on choosing a bus technology (Appendix 1)
2. The concepts resilience and robustness to uncertainty is illustrated on a climate adaptation decision problem related to flood management (Appendix 2)

CONCLUSIONS

Progress made and lessons learnt

The overall aim of this project has been met. We have several operational game prototypes and several prototypes for further development.

Gamification adds new perspectives on science communication. A successful game requires the experts to physically work closely with game developers to find the desirable balance between functional and human based designs. The ShinyR apps that have been developed are tailored towards functional design, whereas the unity games are more influenced by human based design.

A prototype can fail for its purpose when it gets late in a testing phase. Therefore there must be room for experimentation.

We saw that an open communication between experts and game developers creates opportunities for added values.

We also learnt that there is a positive response of gamification, at least initially.

Gamification of decision making turned out to be more tricky than expected. A lot of games already include decision making under risk. How to create a game that explains risk aversion and risk taking that is better than roulette?

A game like the ones developed can be used for many things:

- The primary use of the games is to startle curiosity and motivation.
- A secondary use of the games for expressing your uncertainty and understanding uncertainty in an outcome is for training of experts and testing the skills of experts.
- A third use is to use the game for elicitation of belief and values in decision making problems.
- A fourth use is to let the games produce data for research.

Identified needs

We have had meetings with experts and external partners to identify further needs for this type of outputs and collaboration for additional funding.

We want to continue with the development of games like this, in particular decision making under uncertainty games with links to climate change decisions and the various types of uncertainty therein.

Development of existing prototypes

There is a need for further testing and development of the existing game prototypes. The possibility to get feedback can be further elaborated. Testing on different types of groups will help us to design the games to increase learning even better.

We would like to improve the ShinyR apps and explore ways to use ShinyR or similar softwares for gamification. There is an advantage with using a software which scientific experts are using in their work.

The case-studies needs further improvement. There were no time to test the case-studies within this short project. It was not possible to produce a ShinyR app for the climate adaptation case-study which included probability bound analysis.

Development of new operational game prototypes.

We would like to develop games to motivate to learn more about decision making under uncertainty. Such games links to the task to express uncertainty, since it brings deeper insights into what can be done with what different ways to describe uncertainty.

One idea is to develop a template using the games for training for specific goals that are needed in risk analyses and decision making (both descriptive and prescriptive). This would be a larger project and the results of it could add value to the risk community.

MCDA and serious games can be combined to gain knowledge on player, including assessing player values. Some games could be devised to elicit the values of the subjects for example used in surveys and interviews.

We currently lack a game including a probabilistic characterization of uncertainty for a continuous quantity. It would be possible to fit distributions to judgement on moments, quantiles, tertiles, or densities using the methods e.g. within the SHELF R-package.

It would be useful to create more games connecting probability judgement to a knowledge base and train the gamers to substantiate their judgements. Games could target low probability events or extremes. Learning could be combined with decision making in a game of passive or active adaptive management (Chadès et al. 2017) involving Bayesian learning (Lindley 2006, Cox 2012b). Taking it one step further would be to develop games on learning and decision making under uncertainty linking to robust management and learning under deep uncertainty (Hamarat, Kwakkel, and Pruyt 2013, Cox 2012a).

We have identified a high need in games focusing on creating an interest for forward propagation of uncertainty in models using Monte Carlo simulation, where uncertainty is quantified by frequency and subjective probability or forward propagation using Probability Bound Analysis, where uncertainty using bounds on probability (Ferson and Ginzburg 1996). This could be material for coming workshops and tutorials.

More games on qualitative aspects of uncertainty e.g. as the game aiming to gain an interest in the consequences of framing would be useful.

Final remark

To summarise our conclusions:

- This was fun.
- This is not the end of gamification for better decisions under uncertainty.

REFERENCES

Bhave, Ajay Gajanan, Declan Conway, Suraje Dessai, and David A. Stainforth. 2016. "Barriers and opportunities for robust decision making approaches to support climate change adaptation in the developing world." *Climate Risk Management* 14:1-10. doi: 10.1016/j.crm.2016.09.004.

Burgman, Mark A. 2015. *Trusting Judgements: How to Get the Best out of Experts*. Cambridge: Cambridge University Press.

Chadès, Iadine, Sam Nicol, Tracy M. Rout, Martin Péron, Yann Dujardin, Jean-Baptiste Pichancourt, Alan Hastings, and Cindy E. Hauser. 2017. "Optimization methods to solve adaptive management problems." *Theoretical Ecology* 10 (1):1-20. doi: 10.1007/s12080-016-0313-0.

Cox, D. R. 2006. *Principles of statistical inference*. Cambridge ; New York: Cambridge University Press.

Cox, L. A., Jr. 2012a. "Confronting deep uncertainties in risk analysis." *Risk Anal* 32 (10):1607-29. doi: 10.1111/j.1539-6924.2012.01792.x.

Cox, Louis A. 2012b. *Improving risk analysis*, *International series in operations research & management science,*. New York, NY: Springer.

EFSA. 2016. Draft of a Guidance on Uncertainty in EFSA Scientific Assessment. Accessed Jan 30th 2017.

Ferson, S., and L. R. Ginzburg. 1996. "Different methods are needed to propagate ignorance and variability." *Reliability Engineering & System Safety* 54 (2-3):133-144. doi: 10.1016/s0951-8320(96)00071-3.

Fischhoff, Baruch, and Alex L. Davis. 2014. "Communicating scientific uncertainty." *Proceedings of the National Academy of Sciences* 111 (Supplement 4):13664-13671. doi: 10.1073/pnas.1317504111.

Gage, J., and D. J. Spiegelhalter. 2016. *Teaching probability*, *Cambridge Mathematics*. Cambridge: Cambridge University Press.

Gärdenfors, Peter, and Nils-Eric Sahlin. 1988. *Decision, probability, and utility: selected readings*: Cambridge [Cambridgeshire] ; New York : Cambridge University Press, 1988. Bibliographies

Non-fiction.

Gneiting, T., and A. E. Raftery. 2007. "Strictly proper scoring rules, prediction, and estimation." *Journal of the American Statistical Association* 102 (477):359-378. doi: 10.1198/016214506000001437.

Gregory, Robin, Lee Failing, Michael Harstone, Graham Long, Tim McDaniels, and Dan Ohlson. 2012. *Structured decision making: a practical guide to environmental management choices*: John Wiley & Sons.

Hamarat, Caner, Jan H. Kwakkel, and Erik Pruyt. 2013. "Adaptive Robust Design under deep uncertainty." *Technological Forecasting and Social Change* 80 (3):408-418. doi: http://dx.doi.org/10.1016/j.techfore.2012.10.004.

Herman, Jonathan D., Patrick M. Reed, Harrison B. Zeff, and Gregory W. Characklis. 2015. "How Should Robustness Be Defined for Water Systems Planning under Change?" *Journal of Water Resources Planning and Management* 141 (10):04015012. doi: doi:10.1061/(ASCE)WR.1943-5452.0000509.

Klijn, Frans, Michaël van Buuren, and Sabine A. M. van Rooij. 2004. "Flood-risk Management Strategies for an Uncertain Future: Living with Rhine River Floods in The Netherlands?" *AMBIO: A Journal of the Human Environment* 33 (3):141-147. doi: 10.1579/0044-7447-33.3.141.

Lindley, Dennis V. 2006. *Understanding Uncertainty*. Hoboken, NJ: Wiley.

Linkov, I., T. Bridges, F. Creutzig, J. Decker, C. Fox-Lent, W. Kroger, J. H. Lambert, A. Levermann, B. Montreuil, J. Nathwani, R. Nyer, O. Renn, B. Scharte, A. Scheffler, M. Schreurs, and T. Thiel-Clemen. 2014. "Changing the resilience paradigm." *Nature Climate Change* 4 (6):407-409. doi: 10.1038/nclimate2227.

Maturo, A., and F. Setiffi. 2016. "The gamification of risk: how health apps foster self-confidence and why this is not enough." *Health Risk & Society* 17 (7-8):477-494. doi: 10.1080/13698575.2015.1136599.

Maxim, Laura, and Jeroen P. van der Sluijs. 2011. "Quality in environmental science for policy: Assessing uncertainty as a component of policy analysis." *Environmental Science & Policy* 14 (4):482-492. doi: http://dx.doi.org/10.1016/j.envsci.2011.01.003.

Mens, M. J. P., and F. Klijn. 2015. "The added value of system robustness analysis for flood risk management illustrated by a case on the IJssel River." *Natural Hazards and Earth System Sciences* 15 (2):213-223. doi: 10.5194/nhess-15-213-2015.

Moody, Paul, and Casey Brown. 2013. "Robustness indicators for evaluation under climate change: Application to the upper Great Lakes." *Water Resources Research* 49 (6):3576-3588. doi: 10.1002/wrcr.20228.

Morford, Z. H., B. N. Witts, K. J. Killingsworth, and M. P. Alavosius. 2014. "Gamification: The Intersection between Behavior Analysis and Game Design Technologies." *Behavior Analyst* 37 (1):25-40. doi: 10.1007/s40614-014-0006-1.

National Research Council. 2013. *Levees and the National Flood Insurance Program: Improving Policies and Practices*. Washington, DC: The National Academies Press.

O'Hagan, Anthony, Caitlin E. Buck, Alireza Daneshkhah, and Ei. 2006. *Uncertain Judgements: Eliciting Expert Probabilities*: John Wiley & Sons.

Pidgeon, N., and B. Fischhoff. 2011. "The role of social and decision sciences in communicating uncertain climate risks." *Nature Climate Change* 1 (1):35-41. doi: 10.1038/nclimate1080.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane. 2012. "Forecasting with imprecise probabilities." *International Journal of Approximate Reasoning* 53 (8):1248-1261. doi: 10.1016/j.ijar.2012.06.018.

Spiegelhalter, D. J., and H. Riesch. 2011. "Don't know, can't know: embracing deeper uncertainties when analysing risks." *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences* 369 (1956):4730-4750. doi: 10.1098/rsta.2011.0163.

Troffaes, M. C. M. 2007. "Decision making under uncertainty using imprecise probabilities." *International Journal of Approximate Reasoning* 45 (1):17-29. doi: 10.1016/j.ijar.2006.06.001.

Tversky, A., and D. Kahneman. 1974. "Judgment under Uncertainty - Heuristics and Biases." *Science* 185 (4157):1124-1131. doi: DOI 10.1126/science.185.4157.1124.

van der Sluijs, Jeroen. 2005. "Uncertainty as a monster in the science–policy interface: four coping strategies." *Water Science and Technology* 52 (6):87-92.

APPENDIX 1: CASE-STUDY CLIMATE MITIGATION

Choosing bus technology - climate mitigation

This is a typical situation of choosing between new technologies and is taken from the municipal world. Every municipality has a procurement policy, which addresses CSR, environment, climate, social responsibility or another very similar issue that the procurement policy needs to improve or handle. However, when it comes to evaluating offers for services or goods, in most cases the price alone determines which kind of goods or services that are purchased. The price is assessed with greater weight than values such as CO2 emissions, pollution or other factors. If it is a progressive policy, it may be that it is focusing on life cycle analyzes for the product or on total cost of ownership covering both construction costs / purchase price and subsequent operations costs (i.e. energy consumption associated with the use of the product).

A typical decision problem can be simplified as done in Box 1. Structured decision making is a process to specify decision problems and consider multiple factors when choosing between alternatives (Gregory et al. 2012). Multi Criteria Decision Analysis is a discipline providing theory when multiple factors conflict with each other.

Box 1. What makes a decision?

- An agent
- Her values
- An idea of what is a good decision
- Decision alternatives
- Uncertainties in the outcomes of these alternatives

The fictional game scenario is the problem of choosing a new technology for the greening of the public bus service. In this problem the municipality (**the agent**) must balance factors such as pollution, noise, CO2, and energy consumption with economical values. Here the price is an important factor, but it is equally important to include the subsequent operation (examples of **her values**). Some city councils have even set goals for climate mitigation, but will these goals out beat the effect of cost in the decision process? Take an example of gasoline vs electric cars. It is well known that the car dealer preferably sells a gasoline car because he is sure to see it at the workshop afterwards for service and repairs. While he is very likely not to see the electric car very often. Therefore, they have also introduced regular routine checkups of the battery as part of the prerequisite for the battery warranty, thereby ensuring a certain earnings on the service. On the other hand the buyer may very well like to request the electric car even though it is more expensive to purchase but because it is less costly to run.

Similarly, there may be other criteria for bus operation that are important to consider when choosing a new bus technology for public bus operation. These criteria may be: Total Cost of Ownership (TCO), noise, air pollution, CO2, shaking, purchase price, operating price, infrastructure price, expected cost of maintenance (criteria for which the decision maker has **values**).

In recent times, a handful of technologies has proven market ready for operating public bus services. The conventional diesel bus has been improved in connection with increasing demands for reduced environmental impact when used. Thus, the latest technology is subject to EU standard 6. New technologies include, electricity, hydrogen, compressed natural gas (CNG) and biogas (**decision alternatives**).

Within the electricity segment there are several forms. The traditional ones get driving power from overhead lines. But in recent times it has also been possible to run on batteries. Within battery-powered buses, there are especially two segments, namely those with battery capacity for an entire day's drive and those that needs intermediate charging during the day drive.

Like the electric buses, hydrogen buses have evolved, which in essential is similar to an electric bus with its own power plant instead of a battery. They run with a fuel cell that can produce the necessary power from hydrogen. In addition, there are gas buses that can run on CNG or upgraded biogas. There are a number of other technologies and additional segments to the above mentioned, but these are not included in this case.

One of the major challenges of making a choice between these criteria is that for several of them it is hard to determine a true and exact values (**uncertainty in the outcomes**). If we set up a thought example for some of the technologies, we can see that even if we can set exact values on the criteria, it can be relatively difficult to make a choice between the three technologies.

Traditionally, we begin to figure out what it costs. But how do we value the economic benefits of a CO2 reduction if this should be compared to the purchase costs? And what about comfort how to value the economic consequences of that? We therefore have a challenge and often end up applying a relatively higher weight to the cost criteria than to the other criteria. While we may have relatively exact values for the different criteria, we may not know the cost implications of it, or it may even not make sense to try to evaluate the cost implications for each criteria. Hence it complicates the process of adding a value for each criteria and the subsequent balancing of values of the different criteria so that a final decision can be made. The more criteria that enter the process, the more complex and difficult it becomes. Again, we try to simplify the equation, making it easier to make the decision. This can be done, for example, by collecting some of the criteria in common pools, for example, all criteria for noise, air pollution, CO2, etc. can be collected in a common pool of environment-related criteria. We are therefore trying to simplify the assessment, which of course means loss of accuracy, etc.

Preparing a case for a City Council is challenged exactly by the above situation. The social workers has to build the case on facts and objective arguments, but it simply is difficult to illuminate a case satisfactorily so that it can form the basis for the political treatment and decision subsequently.

An additional source of uncertainty is the possibility of different future developments. It is expected that in the future there will be increasing demand for biogas from the energy sector, which can increase pressure on demand and increased price. Environmental impacts such as Diesel and Gas include a number of social externalities costs, illness, astma etc. Electricity on demand charging requires charging points midway on the routes, but in order to keep costs down, charging points must serve as many bus routes as possible (high utilization rate).

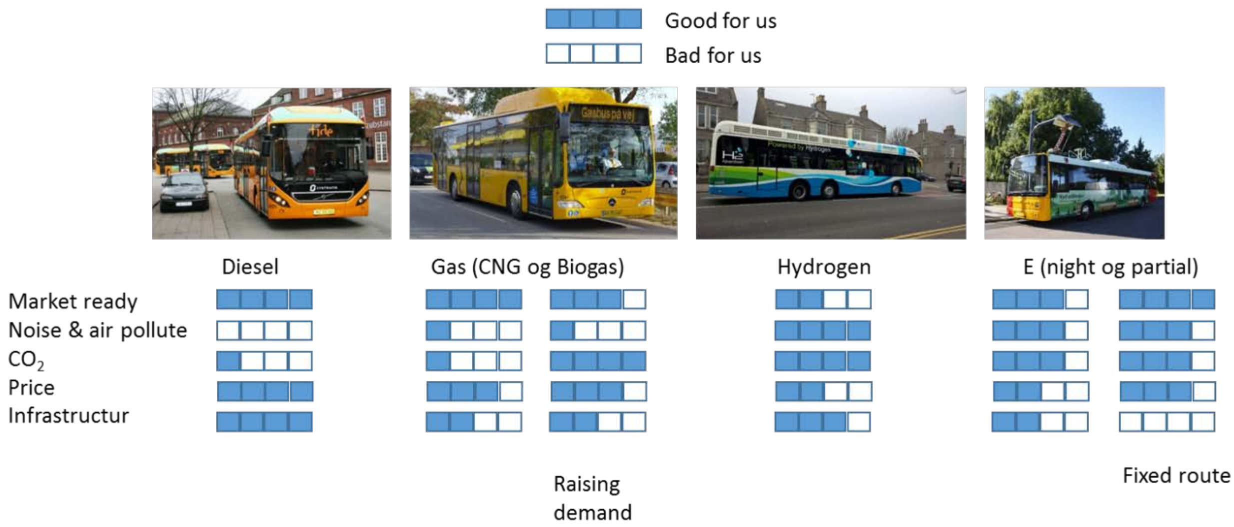
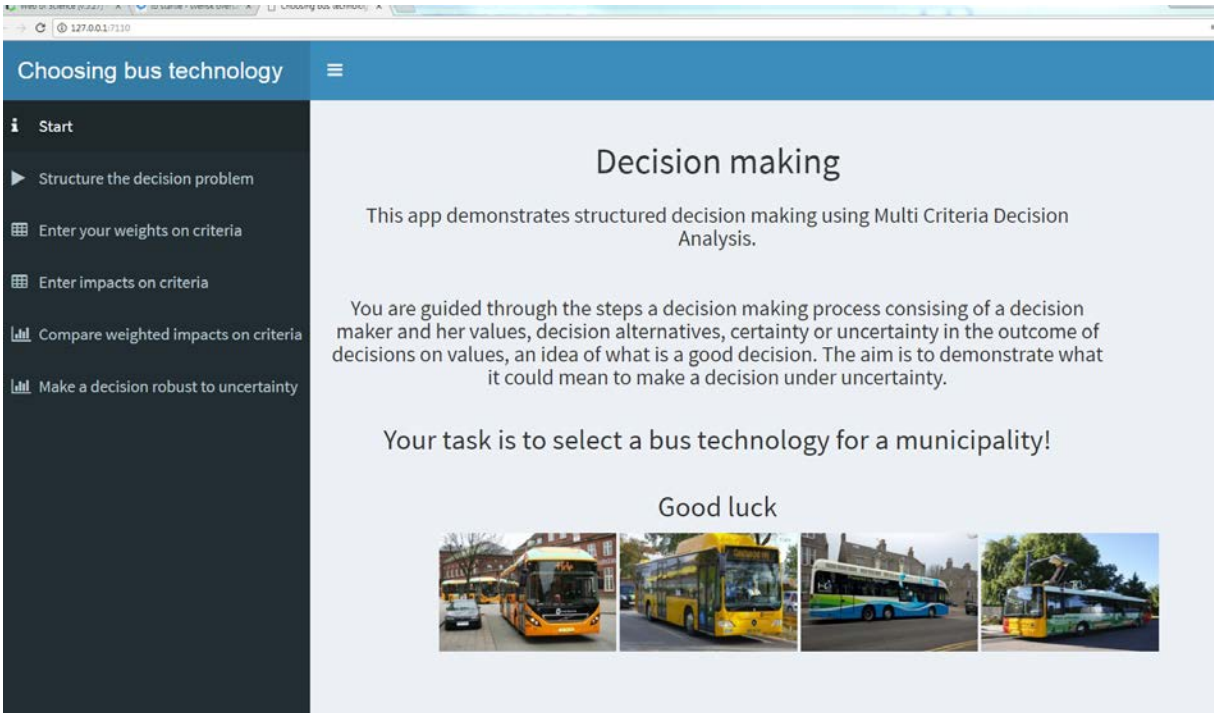


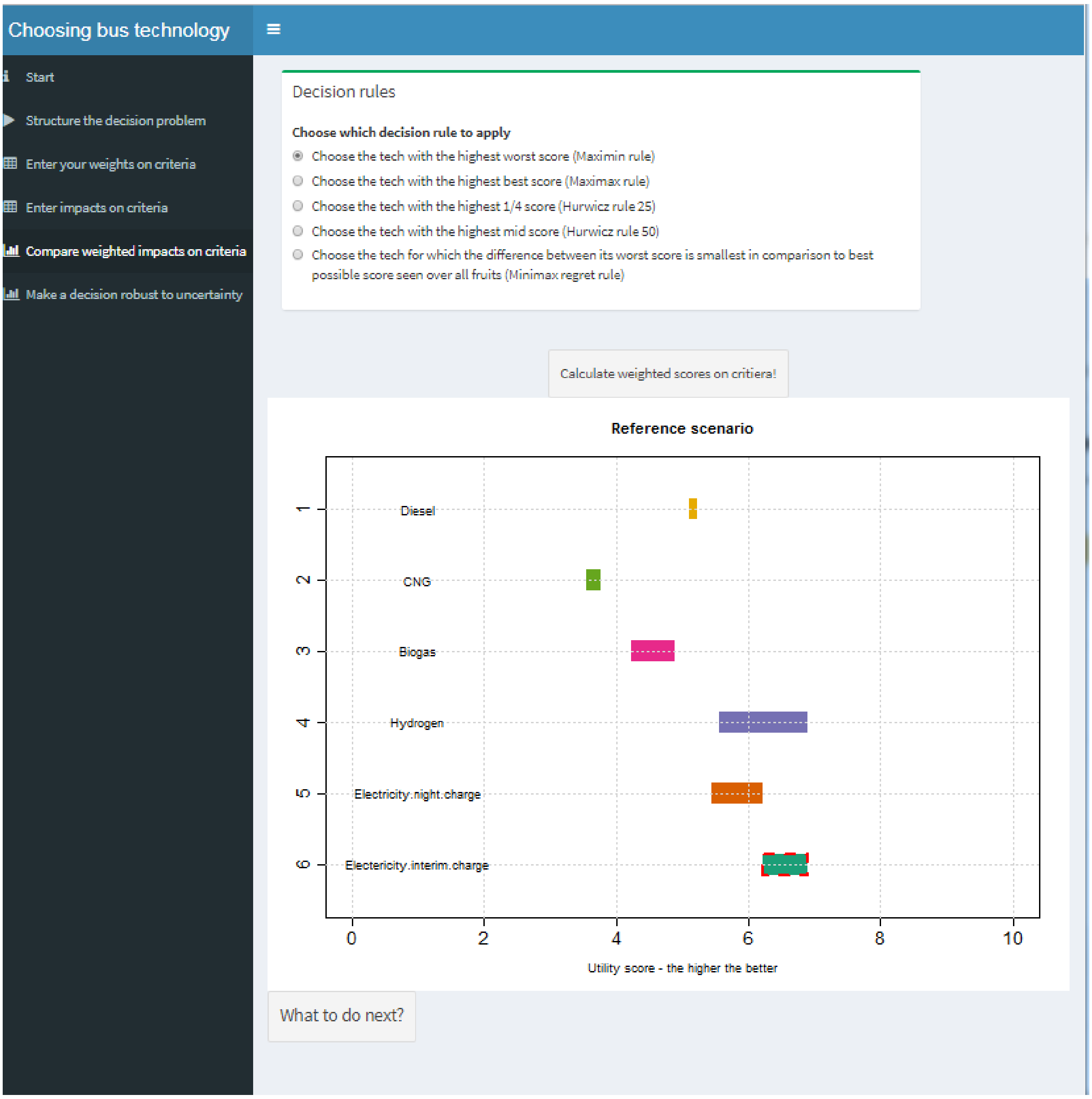
Figure A1.1. Different bus technologies described as a multi criteria decision problem.

Prototype

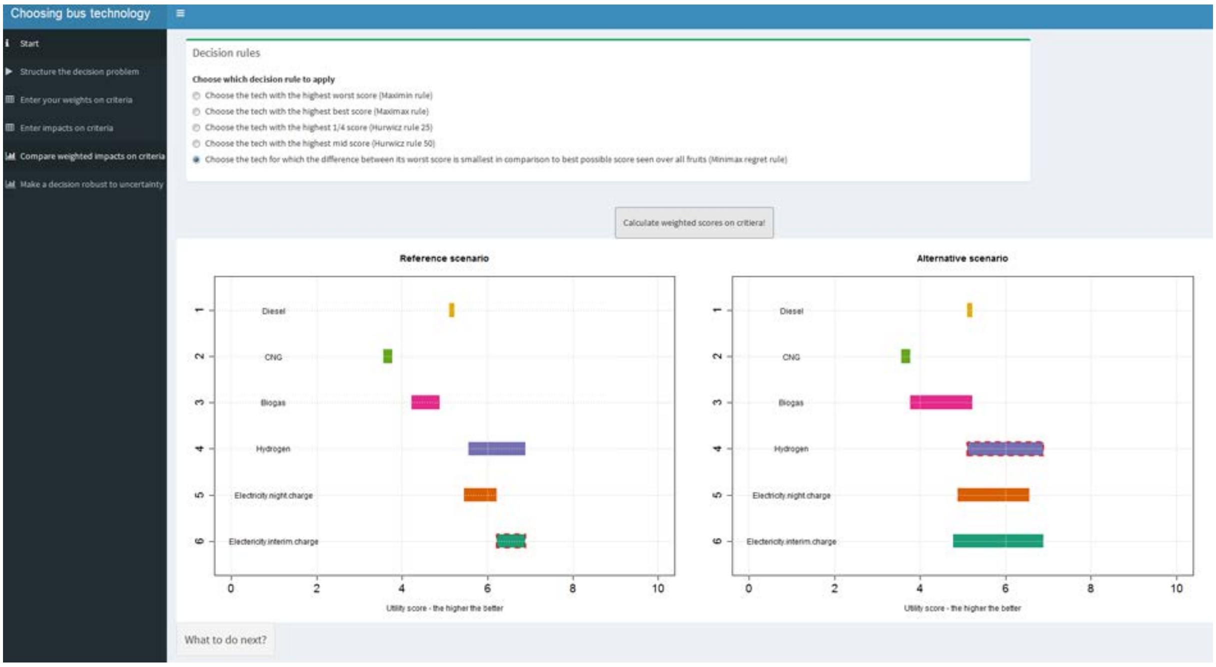
The case-study has been programmed as a walk-through exercise in ShinyR, WhichBusTech.R at https://github.com/Ullrika/GamingForUncertainty



The app goes through a MCDA problem step by step with low amount of technical details. At the end one can explore the preferred alternative under different decision rules. One can go back and see what happens when changing the weights on criteria, scores for the impact on criteria and adding more scenarios.



With more than one scenario it can happen that more than one alternative is the preferred seen over the scenario. In that case- the final tab produces a graph with ranks over scenarios.


Choosing bus technology
Start
Structure the decision problem
Enter your weights on criteria
Enter impacts on criteria
Compare weighted impacts on criteria
Make a decision robust to uncertainty
Decision rules
Choose which decision rule to apply
Choose the tech with the highest worst score (Maximin rule)
Choose the tech with the highest best score (Maximax rule)
Choose the tech with the highest 1/4 score (Hurwicz rule 25)
Choose the tech with the highest mid score (Hurwicz rule 50)
Choose the tech for which the difference between its worst score is smallest in comparison to best possible score seen over all fruits (Minimax regret rule)
Calculate weighted scores on critiera!
Reference scenario
Alternative scenario
Diesel
CNG
Biogas
Hydrogen
Electricity night charge
Electericity interim charge
Utility score - the higher the better
What to do next?

APPENDIX 2: CASE-STUDY CLIMATE ADAPTATION

Flooding resilience and robustness to uncertainty – climate adaptation

From risk to resilience

One way to reduce temporary flooding is to build so called levees, which usually are an earthen embankment, designed and constructed to contain, control, or divert the flow of water so as to provide protection from high water levels. Levees is an example of a strategy that reduce the flooding event to happen, and thus a strategy focusing on reducing a risk. Risk can be defined as the probability of an event, the magnitude in consequences following the event, or a combination of these.

As different from a risk-based strategy, seeking to reduce flooding events, a resilient strategy is to allow temporarily flooding of large areas and adapt land use such that flood damage is reduced (Klijn, van Buuren, and van Rooij 2004). In this case-study, there is a possibility to divert the water to flood a large area with high natural values and thereby reduce the water and likelihood of damage in the urban areas. Flooding the natural area comes with a loss of the high natural values.

Flooding management actions can be structural and non-structural (Table A2.1).

Table A2.1 Examples of Structural and Non-structural Flood Mitigation and Risk Transfer Measures (National Research Council 2013).

Structural	
Levees	Structure elevation
Floodwalls	Natural systems
Seawalls	Risk mapping
Dams	Hazard forecasting, early warning systems, and emergency plans
Floodways and spillways	Dry and wet floodproofing
Channels	Land-use planning and zoning
Controlled overtopping	Construction standards and building codes
Levee armoring	Acquisition and relocation
Seepage control	Insurance

The case-study compare four alternatives to manage flooding:

A. Do nothing
B. Build a spillway to a natural area reducing the risk of flooding in the urban area (“Flood a natural area”)
C. Build a levee to protect urban area from flooding (”Build a wall”)
D. Implement new construction standards and building codes in areas with a high risk of flooding (“Alter houses”)

These four alternatives represent different steps in the resilience framework (Figure A2.2):

Flood a natural area – resilience through risk reduction by absorbing the severity of flood

Build a wall – resilience through risk reduction by preventing flood from occurring, planning/adaptation

Alter buildings – resilience through recovery enhancement by reducing the damages during a flooding event

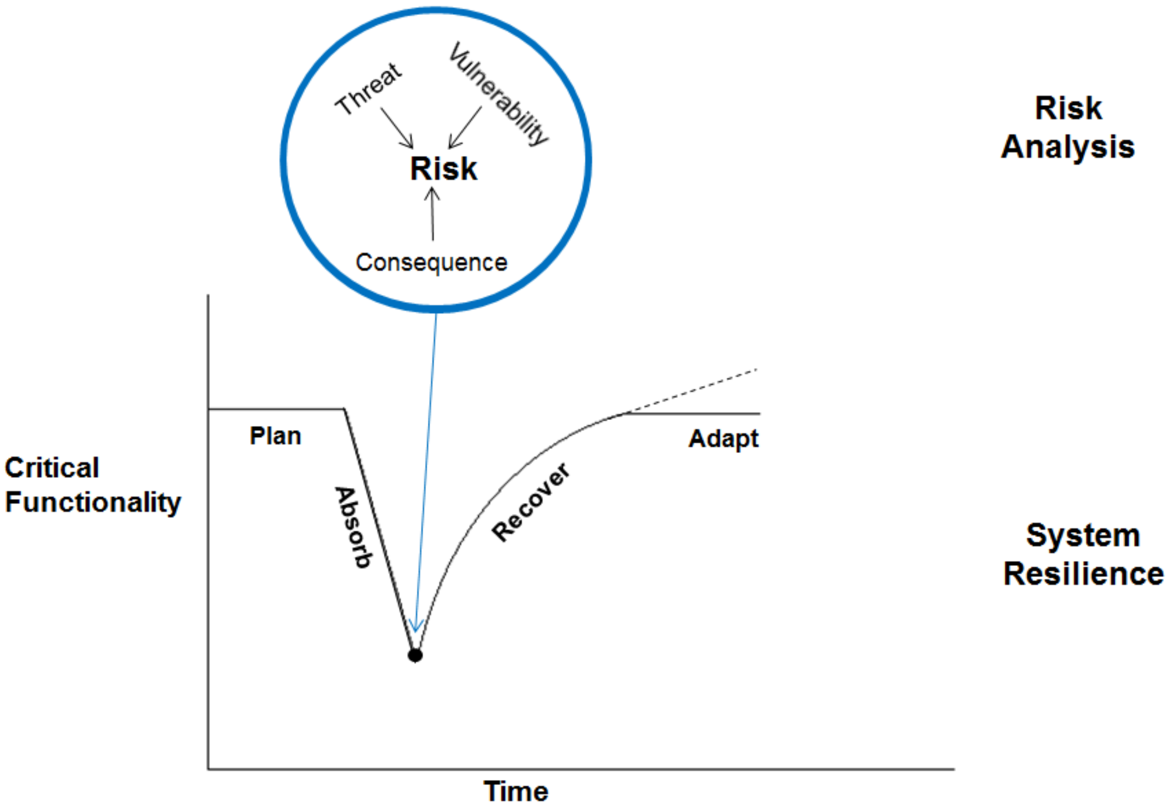


Figure A2.2. An illustration of risk and resilience before and after a hazardous event (from Linkov et al.(2014)).

Robustness to uncertainty

Climate decision problems often face several source to uncertainty. Besides aleatory uncertainty arising from randomness inherent in systems, the ability to predict the outcome of decisions with high precision is reduced by the need to evaluate impacts near and far into the future and by gaps in our knowledge bases (epistemic uncertainty). Sources to such substantive uncertainty (Maxim and van der Sluijs 2011) can be dealt with by uncertainty analysis (Spiegelhalter and Riesch 2011, EFSA 2016) and applying decision strategies coping with uncertainty (Herman et al. 2015, Troffaes 2007). Which decision strategy to use and approach to treat uncertainty in a particular problem depends on the context in which the decision is to be made and the severity of uncertainty in prevailing knowledge. In general, a strategy goes from focusing on characterising uncertainty using quantitative measures to making decisions which are robust to uncertainty, where the impacts from uncertainty on decision objectives are more or less quantified (Cox 2012a, van der Sluijs 2005). Robustness to uncertainty as a performance criteria of decisions made under uncertainty, such as climate mitigation and climate adaptation (Moody and Brown 2013, Mens and Klijn 2015, Bhave et al. 2016).

The case-study was specified to demonstrate resilience and robustness to uncertainty with the purpose to capture relevant features of a climate decision problem. Resilience was included by comparing decision alternatives enhancing resilience in different ways. Robustness to uncertainty was included by including sources to uncertainty of different types, where some were quantified and some were considered as scenarios. Here a scenario is an assumption (or alternative model) for which no weight e.g. in terms of a probability is assigned. Putting a weight on a scenario would result in a multi-model assessment or a hierarchical model, where it can be tempting to take weighted averages over the models when making a decision. Instead, differences between scenarios is an indication of uncertainty in what will happen.

The concept “robustness to uncertainty” is not yet established, and there is several ways to interpret it. First, we conclude that robustness to uncertainty is not the same thing as a robust system, but they are often related. A decision objective can be to have a system robust to hazards and stressors. A robust decision is a decision that is robust to lack of knowledge (i.e. uncertainty). This means that the decision is performing well or good enough even when considering uncertainty in our knowledge about the system. This can be achieved by quantifying and considering uncertainty in an uncertainty analysis or evaluating the decisions across several possible scenarios (scenario analysis). Robustness to uncertainty would also mean that the decision may change when knew knowledge is available, also known as adaptive decision making. Thus, we seek a decision that is robust to uncertainty but sensitive to new knowledge.

Prototype

Cost analysis under the different management alternatives.

Uncertainty in costs and frequency of flooding.

Probability Bound Analysis

Choose decision alternative using robust decision rules.

Introduce scenario.

Choose decision alterative seen over scenario.

ShinyR not possible to combine with the pba code.

APPENDIX 3: SCORING RULES

A scoring rule is a numerical score used to assess the quality of a probabilistic estimation by where the true outcome (event or value) eventually is revealed (Gneiting and Raftery 2007). A scoring rule is proper if the expected score is the highest when the estimation is the same as the true. A rule is proper if the highest value is unique for the true value.

A proper scoring rule motivates the player to give an honest, carefully judged answer, reflecting her uncertainty. Using improper scoring rules creates situations where the player may find strategies which maximise the score but do not give any incentives to give an honest answer.

It is difficult to know how good an estimate of a quantity is when it is only provided by an individual point prediction (i in Figure A3.1). The estimate is either right or wrong. Here we are interested in the player expressing her uncertainty in individual estimates or propositions. This is different from having a set of estimates where one can compare the percentage that are accurate or the 1 to 1 correspondence between estimates/predicted and observed. So, we are interested in scoring rules for individual expressions of uncertainty (e.g. using probability as ii in Figure A3.1). We want to use scoring rules that enhance the user to learn that too precise expressions of uncertainty can miss the true value, while too non-informative expressions of uncertainty are less worth for a possible user.

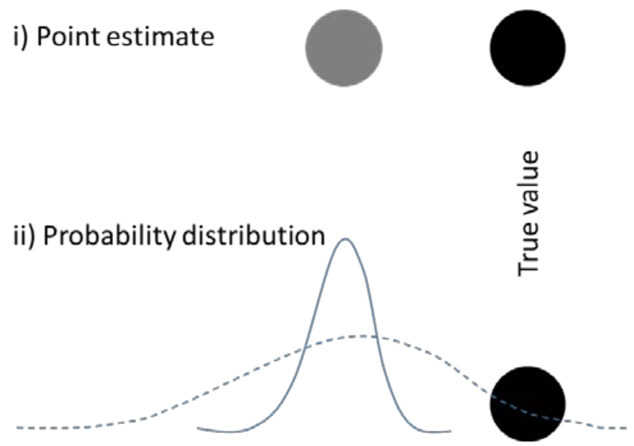
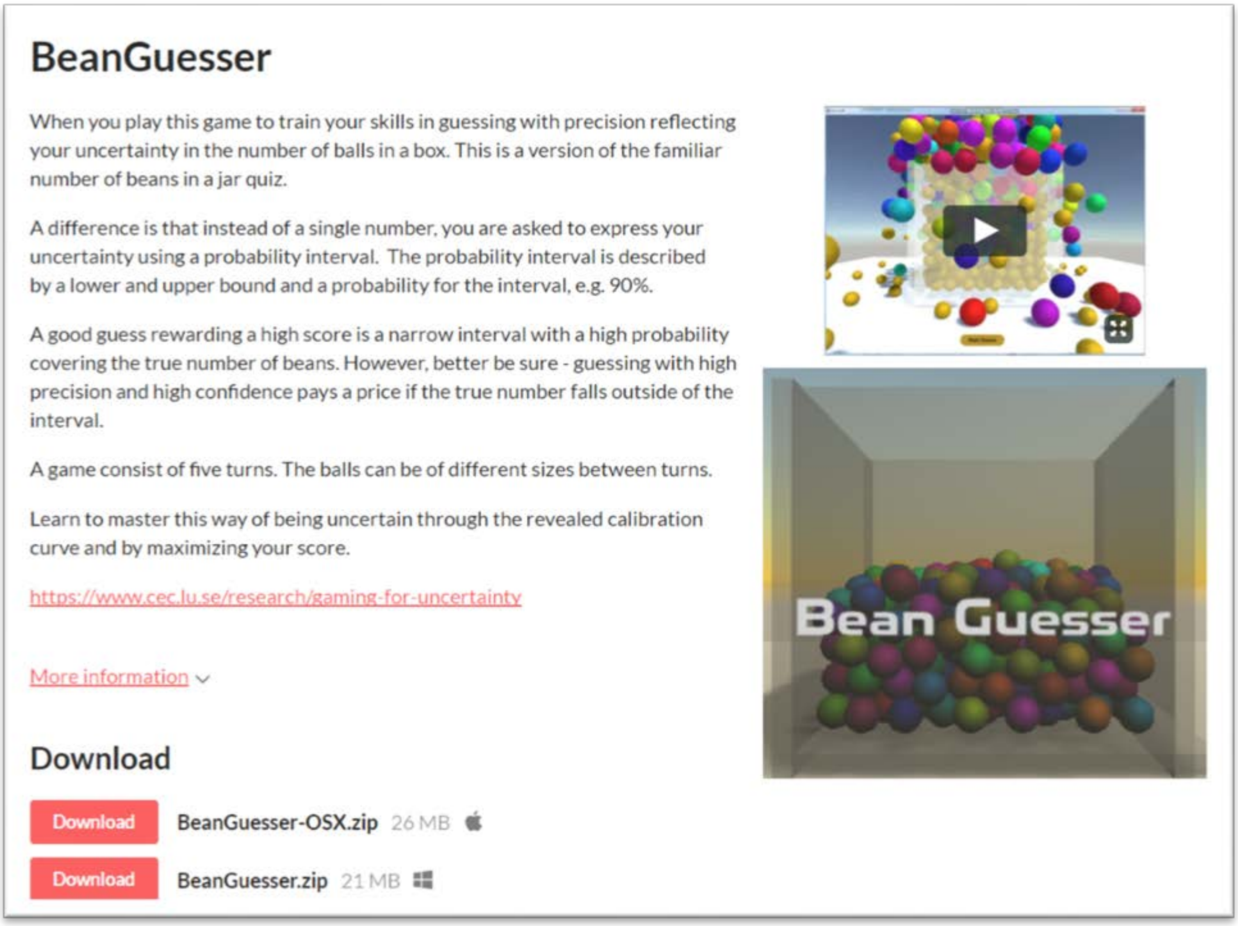


Figure A3.1. An illustration of guessing i) without and ii) with uncertainty, the latter opening up to judge quality in an individual estimate based on the expressed uncertainty and the true value.

Each scoring rule can be scaled to give negative and positive answers which qualitatively tells the player that she is doing a bad and good job.

Feedback is important in order to motivate to continue playing or actually learning something. Besides a score the games were developed to offer some opportunity to get additional feedback.

BeanGuesser


BeanGuesser

When you play this game to train your skills in guessing with precision reflecting your uncertainty in the number of balls in a box. This is a version of the familiar number of beans in a jar quiz.

A difference is that instead of a single number, you are asked to express your uncertainty using a probability interval. The probability interval is described by a lower and upper bound and a probability for the interval, e.g. 90%.

A good guess rewarding a high score is a narrow interval with a high probability covering the true number of beans. However, better be sure - guessing with high precision and high confidence pays a price if the true number falls outside of the interval.

A game consist of five turns. The balls can be of different sizes between turns.

Learn to master this way of being uncertain through the revealed calibration curve and by maximizing your score.

https://www.cec.lu.se/research/gaming-for-uncertainty

More information

Download

Download BeanGuesser-OSX.zip 26 MB

Download BeanGuesser.zip 21 MB

The player assign a probability interval by providing a lower (x) and upper bound (y) and a probability ($p < 0.5$) for the true value being between these two bounds.

For probability intervals there is a scoring rule suggested for quantiles in Eq 43 in Gneiting and Raftery (2007):

$$-\left((y-x)+\frac{2}{\alpha}(x-t)I\{t<x\}+\frac{2}{\alpha}(t-y)I\{t>y\}\right)$$

where $\alpha = 1 - p$. This rule treat the interval bounds as symmetric quantiles. Here the width of the probability interval gives a lower score. If the true value is inside the interval the score is only related to the width of the interval. If the true value is outside, the score is further reduced by the distance from the closest bound to the interval. This is not a proper scoring rule, since the score is not unique for all values inside an interval.

This rule was modified into

$$-100+500\alpha(1-\alpha)-\left((y-x)+\frac{2}{\alpha}(x-t)I\{t<x\}+\frac{2}{\alpha}(t-y)I\{t>y\}\right)$$

A term with α was added to put a penalty for lower probability. Note that α>0. In this case increasing the probability level p would result in a wider interval, creating a trade-off between width and probability (Figure A3.2).

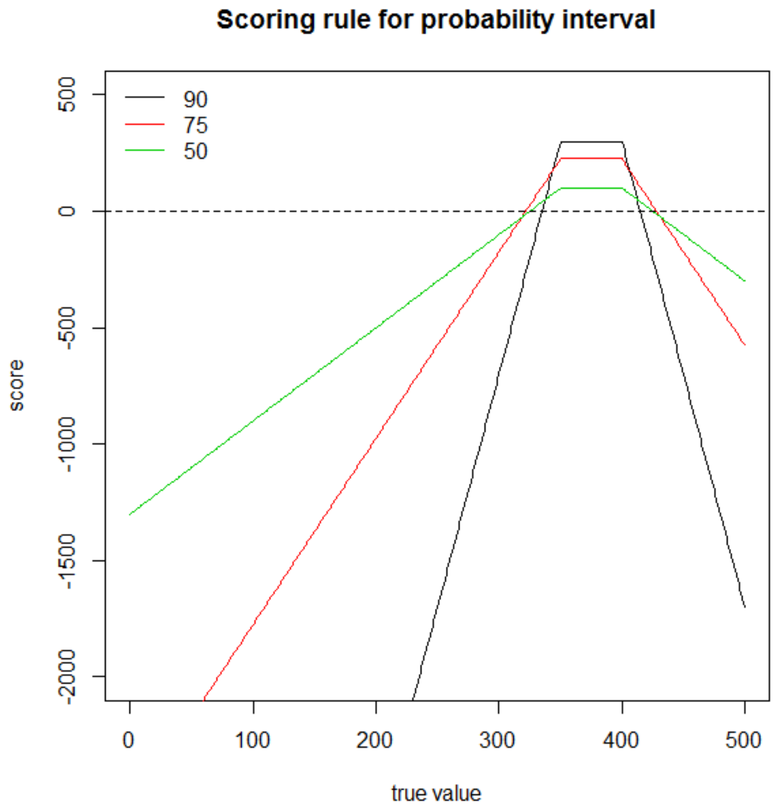
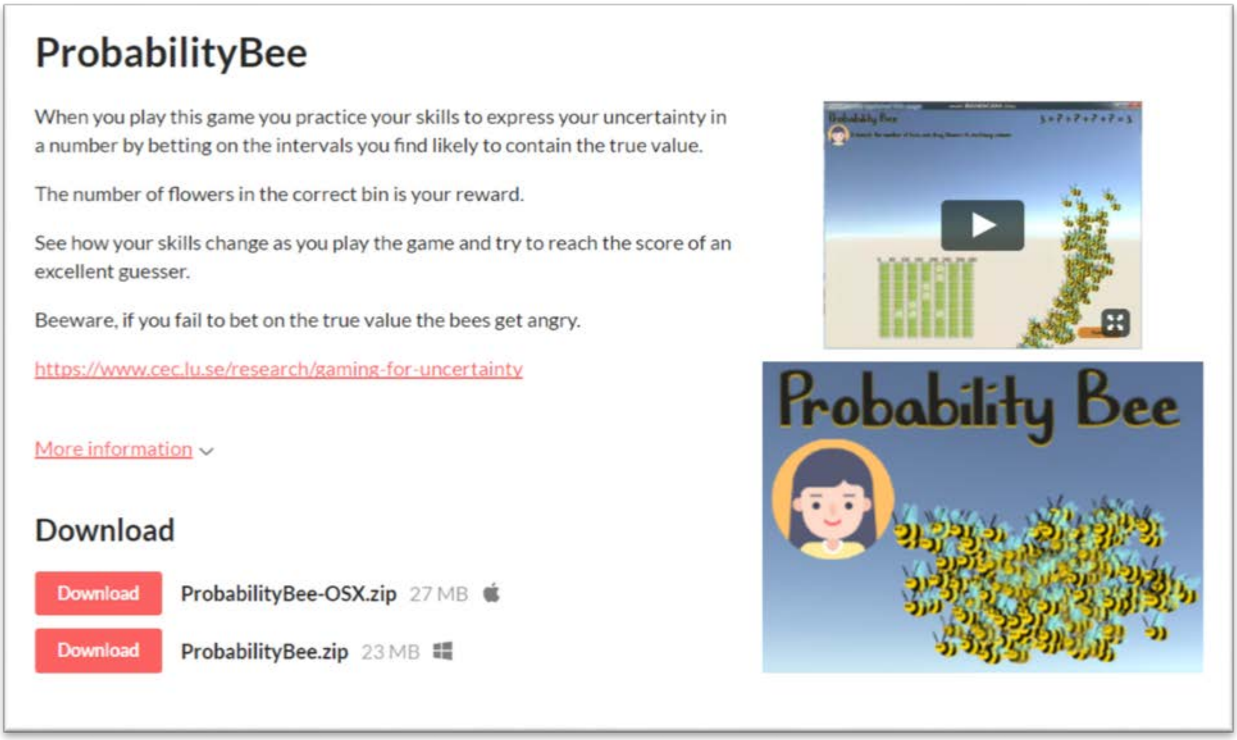


Figure A3.2. The scoring rule for an expression of uncertainty as a probability interval provided by a lower and upper bound (x=350 and y=400) and probability levels in %.

A player get the opportunity to express her uncertainty about the number of beans in five turns. After each turn a comparison on the true value and the probability interval are shown in a graph referred to as the calibration curve.

ProbabilityBee



The player has access to 8 flowers (chips) to place on the field rows (bins) according to her belief in where the true number is. In this way the player bets on the bin and receive the number of flowers put in the bin where the true value is. This is a version of the roulette method where a probability distribution is fitted to have a density matching the chips and bins (Figure A3.3).

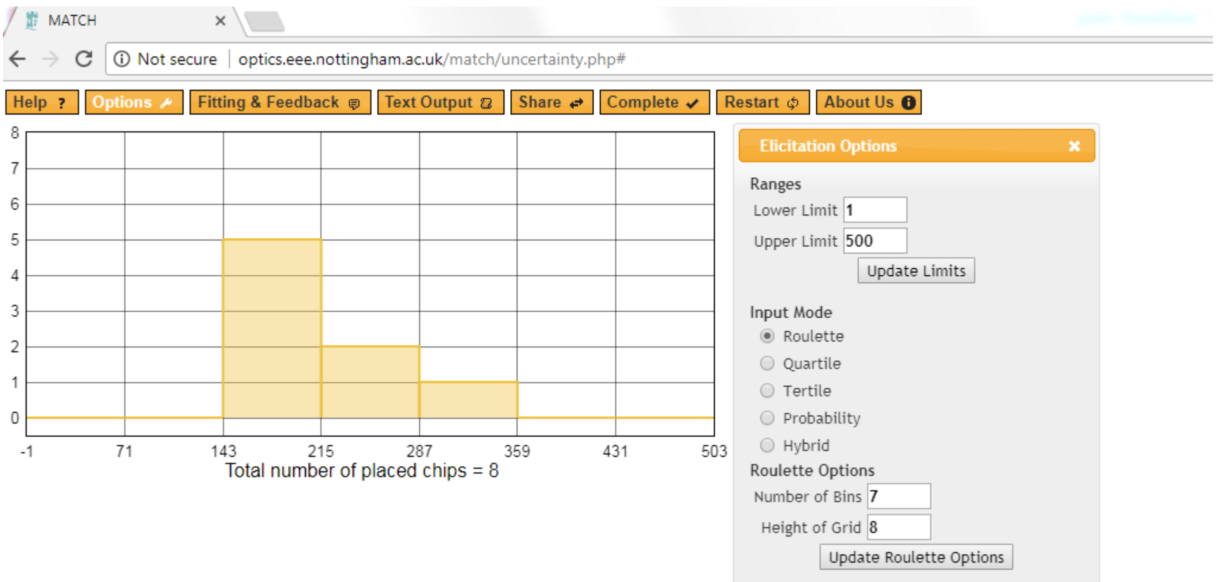


Figure A3.3. A screen saver of the roulette method using the SHELF web version optics.eee.nottingham.ac.uk/match/uncertainty.php#.

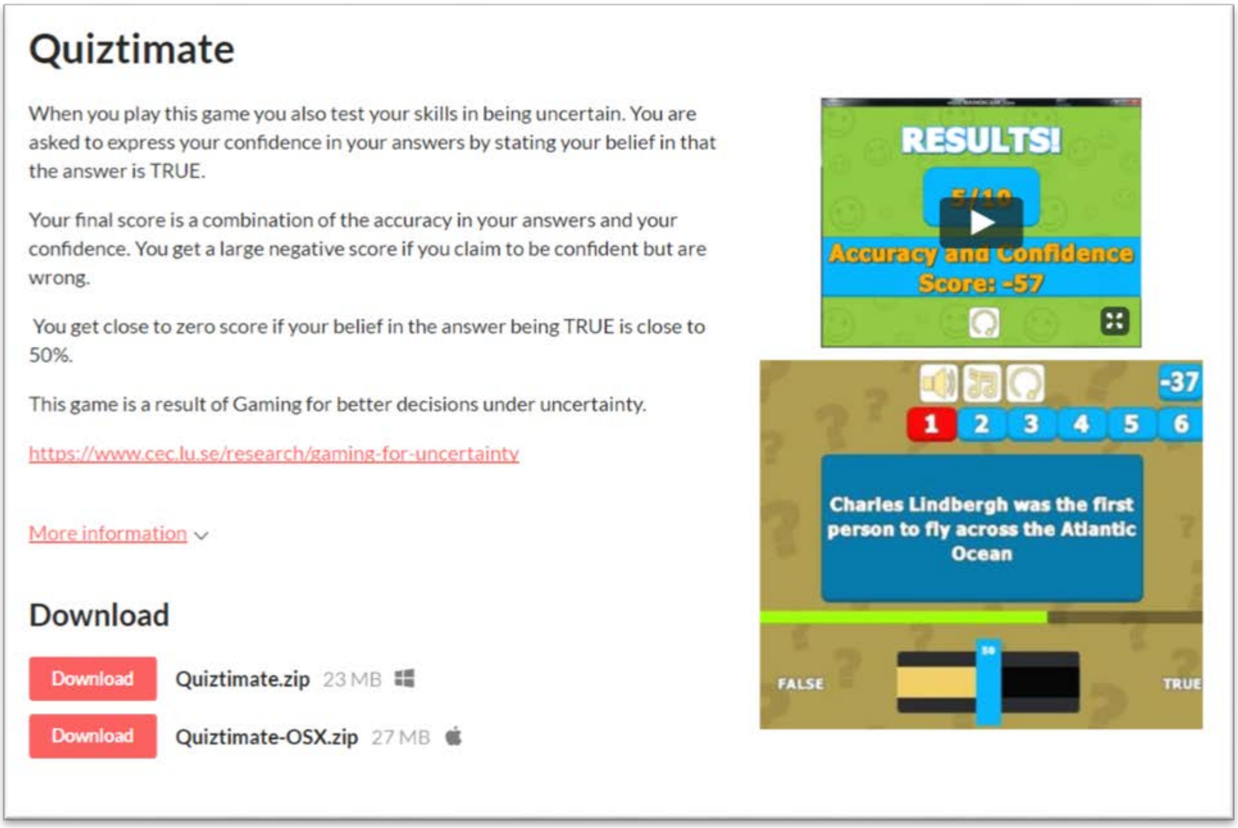
The current version of the scoring rule for ProbabilityBee is not proper but is chosen for its simplicity. It is easy to see how the score is calculated since it is simply the flowers in the accurate bin.

The game is played five turns. Besides seeing the score for every turn, the player also see her learning curve. This is derived by linear regression score ~ turn. A positive curve means that the player improves during the game. In this way one does not only reward the level but also the learning.

A verbal expression of how well the player is doing is shown after the final turn. This expression is based on the total score.

In order to further motivate to make a good guess the bees gets angry when no flowers are placed on the accurate bin (a score of 0). When bees gets angry it can be more difficult to guess the number. Bees become less angry when scores are different from zero again.

Quiztimate



This game is inspired by exercise in chapter 33 in the book Teaching Probability by Gage and Spiegelhalter (Gage and Spiegelhalter 2016).

Uncertainty in a proposition is expressed by a subjective probability, p. The answer to each question is either TRUE or FALSE.

A proper scoring rule is that the score is $-(1-p)^2$ if the answer is TRUE and $-p^2$ if the answer is FALSE (Figure A3.4a). The score in the game is a transformation of this such that the score is 0 when $p = 0.5$, and it takes numbers easy to relate to. The score used in the game (Figure A3.4b) is

$$100 \cdot (0.25 - (1-p)^2 I\{TRUE\} - p^2 I\{FALSE\})$$

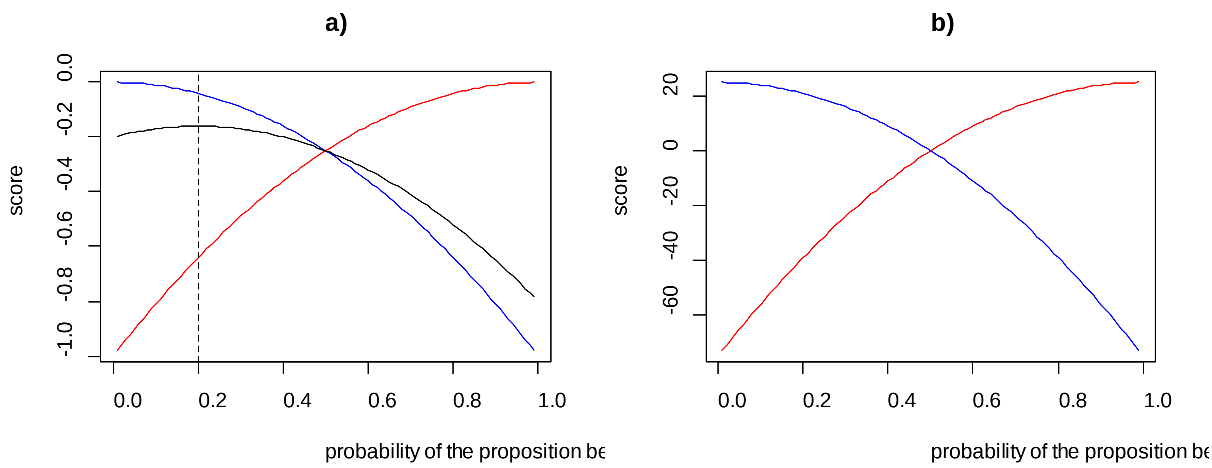
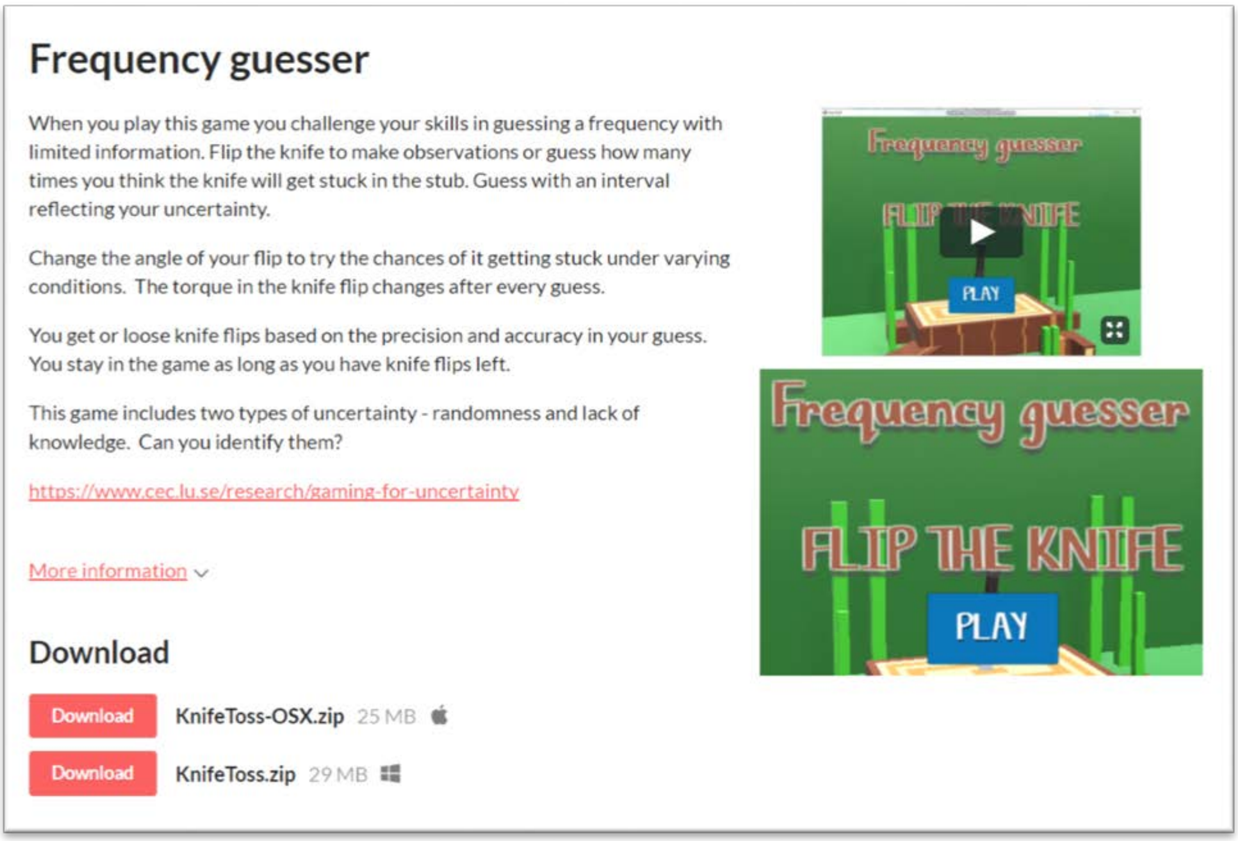


Figure A3.4. Proper scoring rules for guessing a probability seen as the value when the event occurs / proposition is TRUE (red) and when it does not occur / proposition is FALSE (blue). The rule is proper since it takes a unique maximum (black) for the value of the true probability (dashed line in a).

The player see the outcome of each question (ten in total) such that it is green when the player guessed a probability larger than 0.5 on the correct answer and grey if the player put a probability of 0.5. Total score is a sum of the individual scores.

FrequencyGuesser



We decided to express uncertainty about a frequency using an interval. Alternatives include a probability distribution or a probability interval. The frequency is representing aleatory uncertainty, i.e. an inherent random property of the event of the knife to get stuck in the stub. We ask for the number of times out of 100, i.e. a natural frequency, as this is easier to for most people to understand than a percentage (0-100%) or a probability value (0-1). Also, we would like to make the separation between aleatory and epistemic uncertainty implicit.

A proper scoring rule for intervals on probability has been suggested by Seidenfeld et al (2012). This is an imprecise version of the Brier scoring rule and has one score for the lower and one for the upper bound on the interval (Figure A3.5).

Let p be the true relative frequency. The score for the lower bound x is:

$$g = -(1-x)^2 I\{p \geq x\} - (1+x^2) I\{p < x\}$$

The score for the upper bound y is:

$$h = -((1-y)^2 + 1) I\{p \leq y\} - y^2 I\{p > y\}$$

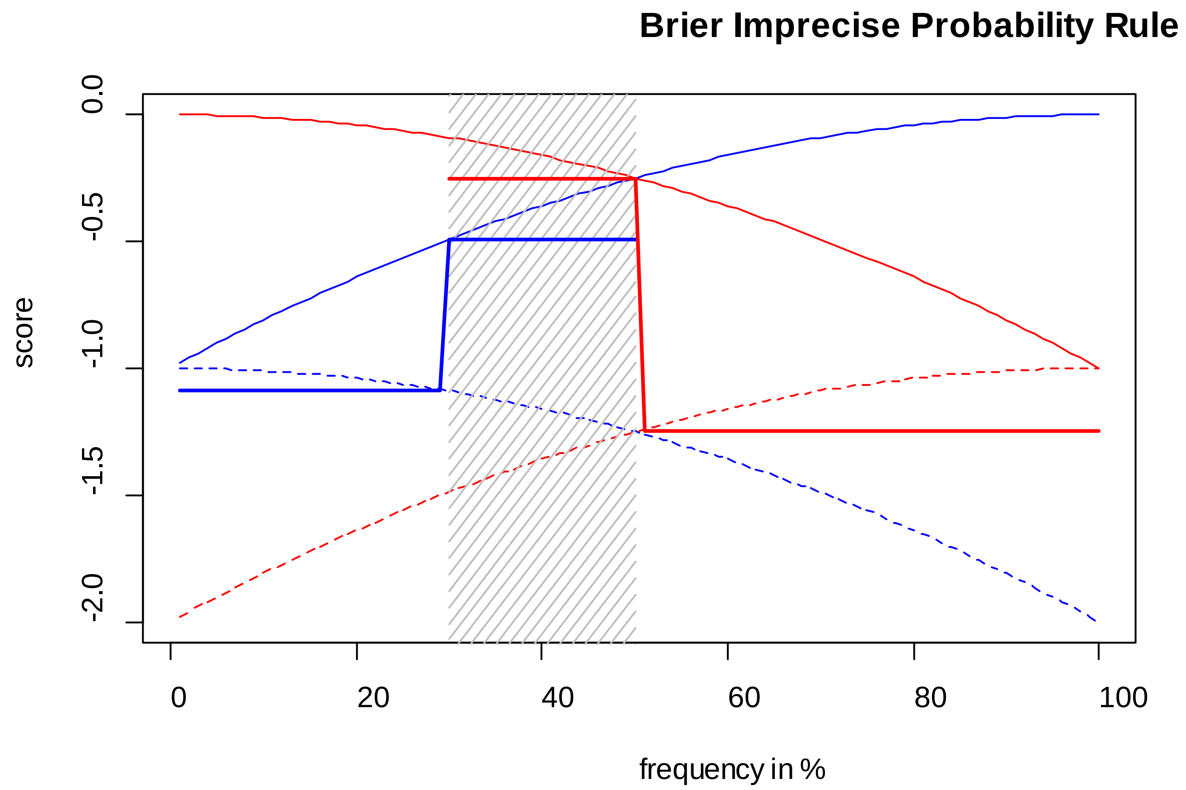


Figure A3.5. Scores for the lower bound (blue lines) and upper bound (red lines), when the true value is inside (solid) and outside (dashed) of the interval.

The scoring rule was modified to give one value by taking the maximum of the score for the lower and upper bounds (Figure A6) and transforming to get score values below and above zero:

$$8 + 16 \cdot \max g, h$$

In this way the player wants to maximise the maximum of the two scores. As can be seen in Figure A3.6 the narrower interval the higher score when the true value is inside and the lower score when the value is outside. This rule is need further development. For example, if the interval is wide enough one can get a positive score even when the true value is outside. How to combine the score for the lower and upper bound is not clear. An alternative would be to maximize the minimal of the score, but this turned out odd in trials.

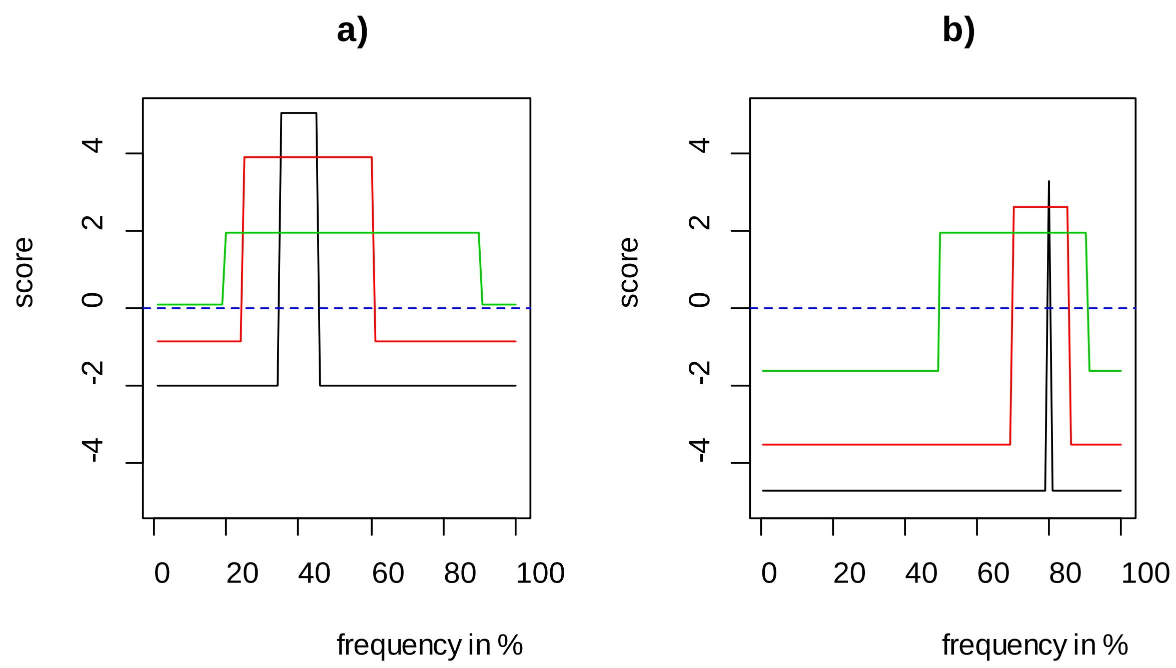


Figure A3.6. The modified Brier score for imprecise probability that is used in the current version of the game shown with three intervals in the centre (a) and bound (b) of the frequency scale. The curves show the score given for the true value on the frequency (x-axis).

The scoring rule either subtracts or adds new flips to what is available to the player. When there are no flips left, the game is over. The player can choose to guess without consuming all flips. The final score is the number of turns the player stayed in the game – put as a level.

The torque (rotation energy) is chosen randomly before every turn. The player can see that the torque changes after every guess. Thus, one should not expect the knife to have the same behaviour between different guesses.

The player can change the angle to test the flip success. In this way, she can do experiments where she collect more observations. The outcome of a flip is identical for a given angle and torque.

APPENDIX 4. TABLE WITH AN OVERVIEW OF GAME PROTOYPES

Table A4.1. Small game prototypes developed to introduce target theory or concepts related to uncertainty and decision making.

Nr	Module	Task	Target theory / concept	Prototype	Program	Link
1	Expressing your uncertainty	Express your uncertainty using probability	Roulette method, betting on outcomes	ProbabilityBee	Unity game	https://uncertaingames.itch.io/probabilitybee https://vimeo.com/248375203
2	Expressing your uncertainty	Express your uncertainty about a quantity using a probability interval	Accuracy and precision	BeanGuesser	Unity game	https://uncertaingames.itch.io/beanguesser
3	Expressing your uncertainty	Express your uncertainty about the frequency of an event using an interval	Aleatory and epistemic uncertainty, frequency and imprecise probability, strength in knowledge, sample size, experimental testing, partial information	FrequencyGuesser	Unity game	https://uncertaingames.itch.io/frequencyguesser https://vimeo.com/248377647
4	Expressing your uncertainty	Express uncertainty in a proposition using probability	Probability as a weight of evidence for a proposition	Quiztimate	Unity ga me	https://uncertaingames.itch.io/quiztimate https://vimeo.com/248368897

5	Understanding uncertainty in an outcome	Test skills in understanding conditional probability vs frequency	Difficulty in understanding conditional probability, communicating aleatory uncertainty. Take part in a historical experiment and compare your outcome with that of the experiment.	In development	ShinyR	
6	Understanding uncertainty in an outcome	Test skills in separating relative frequency from subjective probability	Aleatory and epistemic uncertainty	In development	ShinyR	
7	Understanding uncertainty in an outcome	Explore the meaning of a low probability event	Probability analogies, frequency	LotteryAnalogies	ShinyR	https://github.com/Ullrika/GamingForUncertainty
8	Understanding uncertainty in an outcome	Explore the meaning of uncertainty	Uncertainty analogies, verbal expressions	In development	ShinyR	
9	Decision making under uncertainty	Desirable gambles	Betting interpretation of probability	DecisionMaking1.0	ShinyR	https://github.com/Ullrika/GamingForUncertainty
10	Decision making under uncertainty	Ellsberg's paradox	Uncertainty aversion	DecisionMaking1.0	ShinyR	https://github.com/Ullrika/GamingForUncertainty
11	Decision making under uncertainty	Framing	Cognitive fallacies	DecisionMaking1.0	ShinyR	https://github.com/Ullrika/GamingForUncertainty

12	Decision making under uncertainty	Choosing between alternatives - Multi Criteria Decision Analysis with and without uncertainty	Decision rules, structured decision making, robust decision making, eliciting your values, utility	FruitBreak	ShinyR	https://github.com/Ullrika/GamingForUncertainty