CONDITIONS FOR RISK ASSESSMENT AS A TOPIC FOR PROBABILISTIC EDUCATION

Laura Martignon[1] and Marco Monti[2]
[1]Ludwigsburg University of Education, Germany
[2]Max Planck Institute for Human Development, Germany
martignon@ph-ludwigsburg.de

Statistical literacy is a necessary condition for informed consent and competent citizenship in a modern society. We report data on risk assessment in the context of both medical and investment decision making, which demonstrate that transparent (rather than opaque) communication of risks becomes essential for fostering coherent decisions. We also report on research that reveals how risk communication provided by the pharmaceutical industry or by financial advisors tends to be opaque and excessively complex, rather than transparent and simple. Transparent risk communication can be achieved by making use of frequency statements instead of single-event probabilities, absolute risks instead of relative risks and, in general, "natural" representations of conditional probabilities. We also propose methodologies for instructing children in risk assessments, at an early stadium. These methodologies are based on simple examples.

INTRODUCTION

The neoclassical model of human decision making posited that humans try to maximize their utilities and optimize their use of available information. The flaws of human actual decisions with respect to the neoclassical model were usually explained by cognitive psychologists in terms of "cognitive biases", which prevent humans from being optimizers. This pessimistic view that sees humans as too biased to act according to the norms of logic, probability and utility maximization, has recently been replaced by the more humble paradigm of "bounded rationality" (Simon, 1982). This paradigm views humans not as optimizers but only "satisficers", a term coined by Herbert Simon, who melted the verbs "to suffice" and "to satisfy" into a new verb, namely "to satisfice". Humans, according to Simon, try to find good enough solutions, guided by aspiration levels, which can be quite modest, when information search is costly and deliberation time is limited. Gerd Gigerenzer and his Center for Adaptive Behaviour and Cognition have adopted Herbert Simon`s paradigm of bounded rationality and have gone one step further by adding to Simon's paradigm the framing factor of mind's adaptation to the environment. This factor has been called "ecological rationality". Ecological rationality is rationality within the conditions imposed by the environment, typically by the way information is provided by the environment and by the limits imposed by it. There is, in fact, a flurry of recent literature in economics and psychology documenting that decision makers typically do not incorporate all available information on a given risky situation into their decisions, even when that information is statistically valid, non-redundant (i.e., non-collinear with other predictors), and costless to acquire (Balzer, Doherty, & O'Connor, 1989; Chewning & Harrell, 1990; Lee & Lee, 2004; Berg & Hoffrage, 2008). There is also a plethora of papers confirming Gigerenzer's claims on the mind's ecological rationality, exhibiting how information formats can shape its good or bad interpretation and thus have an influence on decision making. It is fair to say, that Simon's and Gigerenzer's views keep having an impact on recent developments in decision-making and in economic models.

In this paper we discuss a type of risky decisions that are typical in two domains, namely that of investment choices and of medical decision making in situations when time is limited and high risks are at stake. There is one important common feature to these two domains, namely that the decision processes can be modelled by relatively simple decision trees. These decision trees require the capability of the decision agent of ranking features (i.e., characteristics) of the situations involved, according to a measure of their goodness, called their validity (or "diagnosticity"), which is often only intuitively assessed by agents, although it can be formally captured by means of Bayes' Theorem. The alarming findings on base rate neglect when establishing the validity of a positive medical test (say Mammography) for categorizing a patient as having a disease (say Breast Cancer) based on probabilistic formats (the base rate, the sensitivity and the specificity of a test) made popular by the Heuristics and Biases Program of Kahneman, Slovic and Tversky (1982) had

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php [© 2010 ISI/IASE]

long caused a pessimistic attitude among scientists investigating human decision making. Both for the financial and the medical domain, the more recent findings on the benefits of "natural" information formats, like those provided by sheer experience on the samples "naturally" collected by the agents during their own range of experience (Gigerenzer & Hoffrage, 1995; Gigerenzer, 2002) allow for a more optimistic view on human decision making. Humans, according to these findings, are not hopelessly biased when they have to make probabilistic judgments: they are rather dependent on how information is provided to them by their environment. These findings motivate the effort to improve school statistical instruction aiming at familiarizing youngsters with simple instruments for solving tasks similar like assessing the validity of a test, in the case of a disease, or of an investment feature, in the financial realm.

THE ROLE OF CUE VALIDITIES AND HOW TO ELIMINATE BASE RATE NEGLECT

Decisions on actions to be carried out under risk, like choosing an investment or a medical treatment can often be modelled by so called decision trees. As an example, we illustrate a decision tree for investment decisions, which was validated empirically by Monti, Martignon, Gigerenzer & Berg (2009) in the financial domain. The following diagram represents a tree typically used by mature investors (age > 50), when they have to make decisions based on features like risk (high or low), liquidity, duration, cost, etc.

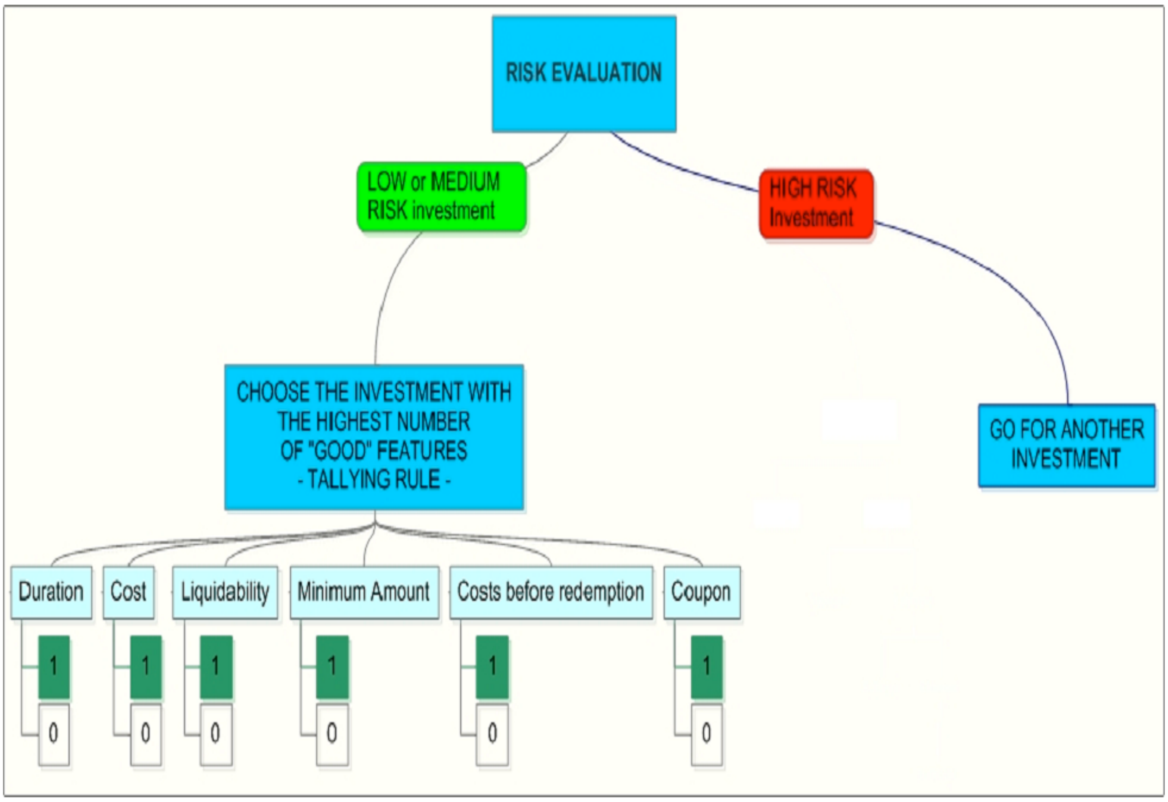

Figure 1. Decision Tree for Investment Decisions

Figure 1 illustrates the typical decision between two investments, when information on features of both investments is provided. This tree exhibits the paramount relevance of the feature "risk" for mature investors. Risk is consulted first, and if it is high, the investor goes for another investment.

The lower part of the tree exhibits a "tallying" procedure, which arises when investors are unable to further rank other features, like duration, durability or costs before redemption. It has been pointed out that the inability to rank the validities (i.e., how good these features are for the choices to be made) is a drawback for the investors (Monti, Martignon, Gigerenzer & Berg, 2009).

Let us now change the domain and look at decision processes in the medical realm. Doctors having to make the decision on whether to send a patient with intense chest pain to the

Coronary Care Unit or to a nursing bed, look at features provided by the electrocardiogram or at typical features of the patient's history. Some aspects of the electrocardiogram are crucial for this type of decision, like an elevation of the so called "ST segment". Other factor, like the so called "main" symptoms, are looked at, when the ST segment is *not* elevated. Similarly, doctors having to make the decision, of whether to prescribe antibiotic treatment to young children suffering from community acquired pneumonia or use other treatments, use quite simple decision trees (Fischer et al., 2002). In Figure 2 we illustrate these decision procedures.

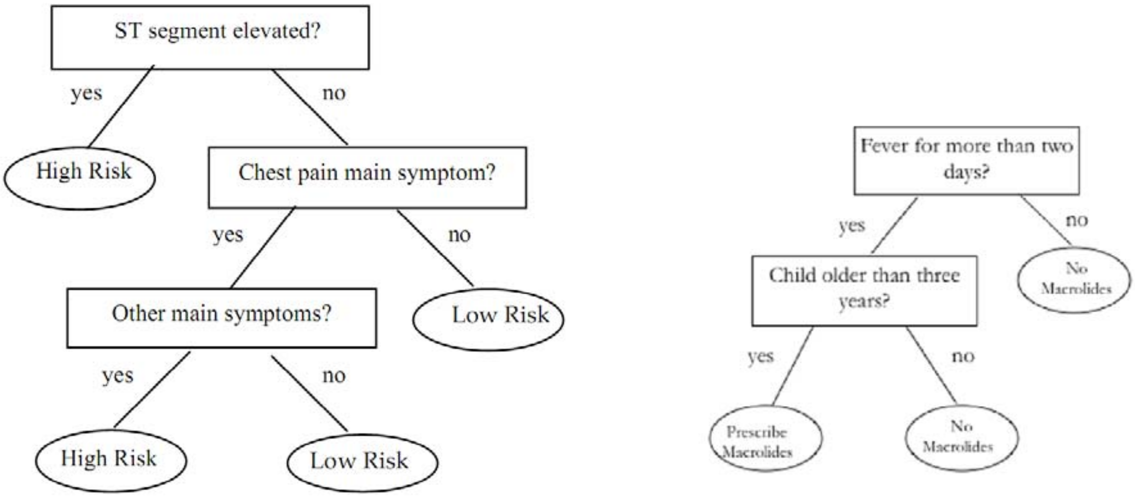

Figure 2. Simple decision trees used in the medical domain

The second tree of Figure 2 first checks whether the child has had fever for more than 2 days or not. If the answer to this is "no", it is immediately concluded that macrolides (the appropriate form of antibiotics) should not be prescribed. Only if the answer to the first question is "yes," the second level of the tree checks whether the child is older than 3 years or not. If the answer to this second question is "no", it is concluded that macrolides should not be prescribed, and if the answer is "yes," then macrolides are recommended. This tree can also be expressed as a rule that is easy to memorize and quick to apply: "*Prescribe macrolides only if the child has had fever for more than 2 days and the child is older than 3 years.*" Additionally, the fast and frugal tree does not sacrifice much in accuracy. When evaluated on real data it classified as "high risk of microstreptococcal pneumonia infection" 72% of those children who actually were at high risk, while a scoring system based on logistic regression identified 75% of them.

We do not illustrate further details of the decision schemes mentioned above but rather point that neither investors nor patients appear to be prepared for formally assessing the *goodness*, or validity of features characterizing treatments or investments. The need for a better communication between experts and their clients or patients, which has crystallized into the concept of „informed consent", has caused an interest in educational measures that may prepare future adults to better understand information on risks and on features for decision making (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz & Woloshin, 2007).

HOW CAN WE INSTRUCT CHILDREN FOR VALIDITY ASSESSMENT

Is the colour "red" a valid feature for identifying a flower as a rose? Is "long hair" a valid feature for classifying a child as a "girl"? What makes a feature valid for a given classification task? Children in fourth grade can understand this question quite well and can analyse the characteristics of good features. In a series of activities with fourth grade children were instructed to encode features by using coloured tinker cubes and forming so called tinker towers.

For instance, they had to encode *boys* as represented, say, by *blue* cubes and *girls* by, say, *red* cubes. *Liking computer games* was encoded by the colour *white* and *not liking computer games* was encoded by the colour *black*. New features could eventually be added: *having or not having*

Math as one's favourite subject. The pair *yellow-green* was introduced to encode this new binary feature (Figure 3 right). The next step was to perform quantified categorizations and introduce the concept of *relative proportions*, which lays the first foundation for the concept of *conditional probability*. The instructor asked, for instance, what proportion of girls likes computer games. (Later in probability theory this can be extended to the probabilistic question: Assuming that a pupil is a girl, which is the probability that she likes computer games?) What proportion of those who like computer games are boys? (Assuming that a pupil likes computer games, what is the probability that this is a boy?)

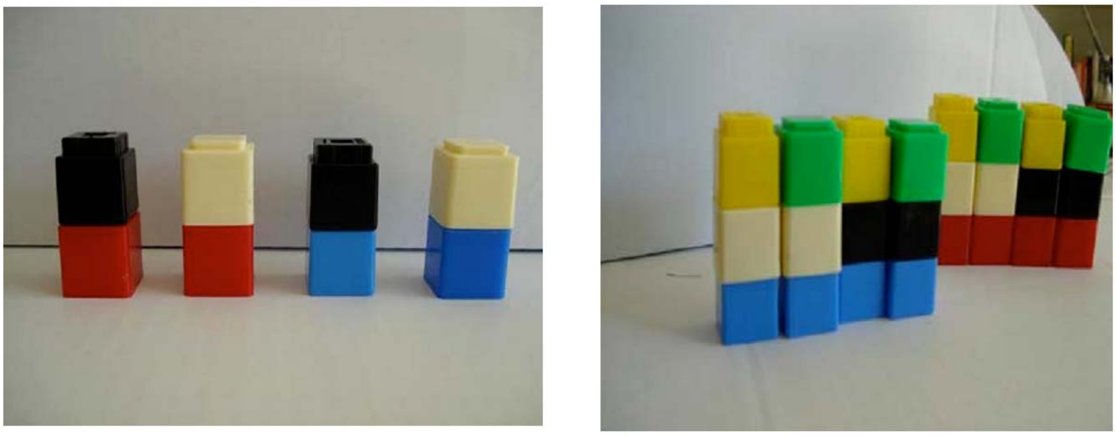

Figure 3. Tinker towers for 2 feature conjunctions (left),
representing 3 features conjunction (right)

The success of these types of activities was evaluated in tests, where children had to model situations similar to those described above and answer questions on the validity of features (Martignon & Krauss, 2009).

RELATIVE AND ABSOLUTE RISKS

The way risks are communicated to the decision makers plays a paramount role on how they make risky decisions both in the context of investments and in the medical domain.
Consider the following real example from an advertisement in the web (see Doctor's Guide) from http://www.pslgroup.com/dg/7DEB6.htm): *Heart-Attack Risk Reduced By Almost Half with Aggrastat.*

This type of information is very frequent, if not typical, in both medical and financial advertisements and brochures. But what does "by almost half" actually mean? How should patients understand the risk reduction achieved by using Aggrastat?

Let us consider another instance of the same phenomenon (Gigerenzer, 2002): Women are constantly prompted to perform regular screening (mammography) to enhance their life expectancy of 25%. The 25% enhancement of life expectancy can be explained as follows: Four out of 100 women have breast cancer, and if these 100 women perform screening regularly then one will be saved. It is crucial to become aware that the corresponding percentage would be 0,01% (1/100) when communicated as absolute risk.

Let us summarize some typical forms of risk communication:

- *Relative risk reduction*: Mammography screening reduces the risk of dying from breast cancer by 25 percent.
- *Absolute risk reduction*: Mammography screening reduces the number of women who die of breast cancer by 1 out of 100.
- *Number needed to screen*: In order to prevent one death, the number of women who need to participate in screening is 100.

INSTRUCTING YOUNG CHILDREN IN RISK ASSESSMENT

What is risk? Risk analysts have given complex definitions of this term. For common usage in everyday life it can also be defined quite simply as a characteristic of experiments, in which at least one of the outcomes is coupled with a substantial loss of resources. Risks are usually described in terms of gambles, like the following:

Do you prefer to get 10€ immediately or to throw a coin and get 20€ if it turns "heads" and 0€ if it turns "tails"? Gambles like these appear to be well understood by a good portion of children. This is, at least what findings of an empirical study by Engel, Kuntze, Martignon and Gundlach (RIKO-STAT, 2009) reveal. Children in fourth grade (aged nine to ten years) appear to grasp that the offer is not unfair and that the second option involves "risk". In order to instruct young children on the difference between absolute and relative risks, Engel, Kuntze, Martignon and Gundlach (RIKO-STAT, 2009) have devised simple tasks associated to decisions in typical ritualized games, like LUDO, that are played in most western countries.

Consider the following situation in LUDO: there only two players left, namely "Blue" and "Green". Blue is two squares behind Green and at another position Blue is three squares behind Green. Assume it is Green's turn. He draws a "one". Which token should Green move? Which move is riskier? The risk of losing one token can have a probability of "one out of three" if one moves further with the right token (Figure 4) or "one out of six" if one moves with the left token.

How do we compare the risk in the two scenarios? Would it be correct to say, that *"the expectancy of Green keeping a token has been increased by 100%"*, when actually *the risk of losing one token has been reduced from "two out of six" to "one out of six"?*

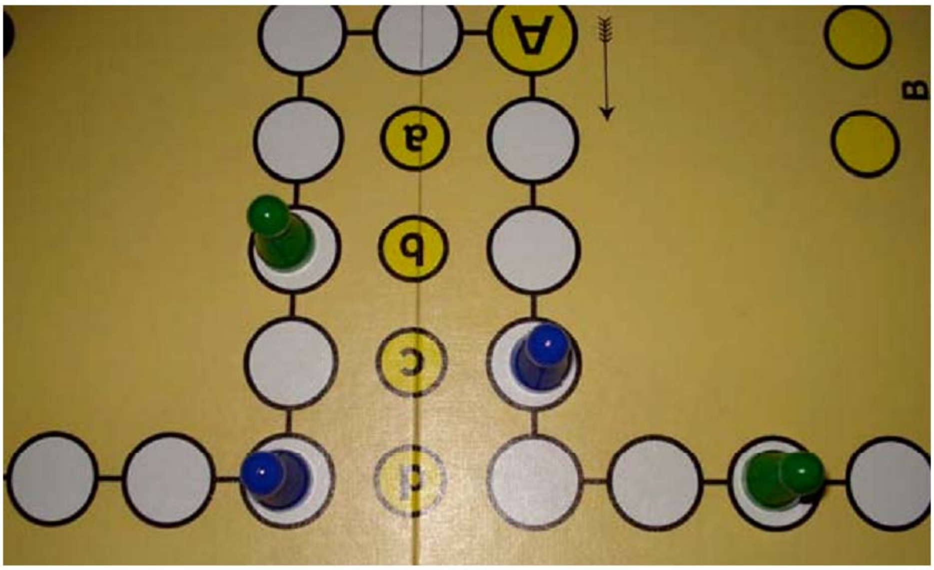

Figure 4. A typical setting of the LUDO game: Green has drawn a "one" with a die
Which move should Green make to reduce the risk of losing a token?

The test results with four graders were encouraging as far as the understanding of risk is concerned. Moreover, with older children in ninth grade, we had the positive experience that not only did they understand the risk involved in the decision to be made by Green but also grasped, that risk can be expressed in two very different ways, which may cause quite different reactions.

DISCUSSION

This paper is aimed at convincing future teachers that it makes great sense to prepare children for an adult life that can deal with risks with the help of simple statistical instruments learned at school. It also aims at presenting possible ways for en-actively instructing children on basic issues about validities of features for decisions under uncertainty and for reckoning with risks. It motivates the instruction on different ways for communicating about risks, so as to make clear that some forms of communication can be transparent while others may be opaque and amenable to unnecessary misunderstandings.

REFERENCES

Batanero, C., Burrill, G., Reading C., & Rossmann, A. (2010). *Joint ICMI/IASE Study: Teaching Statistics in School.*

Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin*, *106*(3), 410-433.

Chewning, E., & Harrell, A. (1990). The effect of information overload on decision makers' cue utilization levels and decision quality in financial distress decision task. *Accounting, Organizations and Society*, *15*(6), 527-542.

Doctor's Guide: http://www.pslgroup.com/dg/7DEB6.htm (retrieved March, 23, 2009)

Gigerenzer, G. (2002): *Reckoning with risk: learning to live with uncertainty*. London: Penguin Books.

Fischer, J., Steiner, F., Zucol, F., Berger, C., & Martignon, L. (2002). Use of simple heuristics to target macrolide prescription in children with community acquired pneumonia. *Archives of Pediatrics and adolescent Medicine*, *156*, 1005-1008.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients to make sense of health statistics. *Psychological Science in the Public Interest*, *8*, 53–96.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102-4,* 684-704.

Kahnaman, D., Slovic, P., & Tversky, A. (1982). *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Lee, B. K., & Lee, W. N. (2004). The effect of information overload on consumer choice quality in an on-line environment. *Psychology & Marketing*, *21*, 159–183.

Martignon, L., & Krauss, S. (2009). Hands-On Activities for Fourth Graders: A Tool Box for Decision-Making and Reckoning with Risk, *International Electronic Journal of Mathematics Education*, *4*(3), 227-258

Monti, M., Martignon, L., & Gigerenzer, G. (2009). Le decisioni nel marketing finanziario:dall'homo oeconomicus all'homo heuristicus (Financial Decision-Making: from Homo Oeconomicus to Homo Heuristicus), chapter N°3. In A. Antonietti & M. Balconi (Eds.) *Valutare, Scegliere, Comprare - Prospettive recenti di psicologia e neuroscienze circa le dinamiche di acquisto* (Marketing and financial decisions: insights from psychology and neurosciences for explaining economic behaviors). Berlin, Edizioni Springer.

Monti, M., Martignon, L., Gigerenzer, G., & Berg, N. (2010). Fast and Frugal Heuristics in High Stakes Investment decisions: New Descriptive Models of Bank Customers' Behavior, (under revision for *Mind and Society*).

RIKO-STAT (2009-2011). Project of the PH Ludwigsburg, directed by J. Engel, S. Kuntze & L. Martignon.

Simon, H. (1982). Models of bounded rationality. Economic analysis and public policy, 1982 - mitpress.mit.edu.