



**International Electronic Journal of
Mathematics Education**

Volume 4, Number 3, October 2009

www.iejme.com

Special issue on “Research and Developments in Probability Education”

Manfred Borovcnik & Ramesh Kapadia (Eds)

**HANDS-ON ACTIVITIES FOR FOURTH GRADERS:
A TOOL BOX FOR DECISION-MAKING AND RECKONING WITH RISK**

Laura Martignon & Stefan Krauss

ABSTRACT. The intention of this work is to exhibit how children can be provided with a kit of elementary tools for judgment under uncertainty, for good decision making and for reckoning with risk. Children, we claim, can acquire this tool kit through a mosaic of simple, play-based activities which are devised to make them aware of the characteristics of uncertainty. We present a sequence of tasks that build upon each other, beginning with the Wason selection task, moving on to probabilistic tasks, tasks in elementary Bayesian reasoning comparing proportions and, finally, to comparing risks. This research is guided and inspired by empirical results on human decision making in the medical and financial domain.

KEYWORDS. Tool kit, Tinker Cubes, Bayesian reasoning, Wason’s experiment, risk.

1. INTRODUCTION

Stochastic literacy is a necessary condition for enlightened decision making in an information-based society. Becoming conscious that judgments about our fellow human beings and about nature should often be based on probabilistic rather than strict logical implications can reduce the impact of prejudices and stereotypes. Moreover, an understanding of probabilities can shape our decisions allowing us to assess possible risks associated with our actions. In fact, good modelling of risky situations can sustain our cognitive and emotional perspective on personal and collective affairs, reducing our anxieties and guiding our informed consent.

Copyright © 2009 by GOKKUSAGI

ISSN: 1306-3030

Our conviction is that the basic ideas of stochastic literacy for judgment and decision-making should be conveyed to children at an early stage, more precisely, before they reach their tenth year of age. This conviction is based on the view of experts, who sustain that the mathematical competencies of adults in general, not especially trained in mathematical subjects, are those they develop in first 5 to 6 years of their education; such a view was shared, for instance, by participants of the *Conference on Strategies for Risk Communication* of the New York Academy of Science (see Peters 2008, Finkel 2008, and Kurz-Milcke, et al 2008).

In other words, while most of the mathematical training of secondary school tends to be soon forgotten by those who do not pursue a career requiring further mathematical tools, mathematical competencies acquired until the ninth or tenth year of age appear to remain robust and unaltered during subsequent life. Another feature of our research is our effort to devise tasks that children can perform through hands-on activities. It is an old wisdom confirmed by research on developmental psychology (see Hirsh-Pasek & Golinkoff 2003) that movement and real activity of hands and body, as well as play with simple toys or with each other, is not just healthy but can also be successfully implemented as a tool to enhance learning. Elementary arithmetic and geometry have traditionally profited from playful activities in the classroom. Statistics and probability offer even greater possibilities for “learning by playing”, because the essence of these disciplines is sampling, sorting, betting, and gambling. The attitude of play-based discovery can be regarded as an intrinsic characteristic of stochastic thinking (Amit & Jan 2007).

That children have probabilistic intuitions is an empirically tested thesis widely illustrated in the pertinent literature. It is beyond the scope of this paper to present an overview of it. We base our research on the approach of Fischbein and the large community that confirmed and developed his discoveries. Probabilistic intuitions do not necessarily imply a full understanding of the modern concept of probability. To grasp why the whole body of modern probability makes sense, requires substantial mathematical expertise and not only cognitive maturity while the fundamental properties of probability seem to link to some of our basic intuitions.

In fact, there is evidence that proves how good fourth graders are at comparing events as to their likelihood, although they often make mistakes assessing the probability of the union of events. Typically when they estimate – in percentages – the chances that the different soccer teams win, say, the Champions League Cup, the order relations between each two of their estimates are adequate although single estimates also add up to much more than 100%. Children have intuitions and can be trained in judgments as to whether a specific event is “more likely” than another, or as to whether a given event is “very likely”, “completely sure” or rather

“unlikely” (Neubert 2007; Martignon & Kurz-Milcke 2006). They also enjoy discovering that coins or dice may be “unfair” (Martignon & Krauss 2007) or that even the rules of a game may be “unfair” (Amit & Jan 2007). The current position, at least in Germany, is that *some* aspects of probability are part of children’s early reasoning but not all, and that it makes sense to foster and stimulate these intuitions during elementary school.

2. DECISION-MAKING: FROM ADULTS TO CHILDREN

Our contribution will be devoted to a different set of tasks than those typically treated in the literature. We will describe play-based yet ritualized activities for fourth-graders that have been conceived for training children in the logical and probabilistic subtleties of judgment and decision-making. The path we follow is guided by the chain of competencies that make up good decision making for informed consent in basic domains of modern human life like those of medical and investment decisions. This chain of competencies makes up a tool box for decision heuristics in the bounded rationality paradigm (Gigerenzer, et al. 1999; Todd, et al 2010; Gigerenzer & Brighton 2009; Martignon, et al 2008 and 2010).

Principles of adult decision-making

Let us briefly illustrate what is meant by bounded rationality. Humans are not full optimizers when making their day-to-day decisions. Their decision making is not Bayesian, which usually requires excessive computational complexity. They rather combine elements of basic Bayesian reasoning with effective decision heuristics. Humans are successful “satisficers”, to use a term typical in the bounded rationality approach: The verb “to satisfice” melts the verbs “to satisfy” and “to suffice” and was used by the Nobel-laureate Herbert Simon, to describe human strategies for decision making. In a broad variety of situations, humans use features to categorize possible alternatives to which they associate possible actions. When more than two features have to be taken into consideration, the resulting complexity makes humans switch to very simple heuristics for handling multiple features.

In fact, humans often proceed by ranking features according to their validity, computed in a Bayesian way, eliciting feature values, one by one, following their ranking, until they reach a decision. Thus doctors often follow decision routines based on sequences of tests, chosen according to their probabilistic validity without integrating test results in a full Bayesian way.

Let us present an example from medical decision-making. The tree we depict in Figure 1 is used to make the decision whether a patient with severe chest pain should be assigned a nursing bed (low risk) or sent to the coronary care unit (high risk). The main features used for this decision are characteristics of the electrocardiogram and statements of the patient on perceived symptoms. Bayesian reasoning is used for computing the validity (or diagnosticity, see the [appendix](#)) of each single feature, that is the probability that the patient is at high risk given a positive feature value. The assessed validities of all features considered are then compared to each other and ranked from highest to lowest. This ranking allows constructing a so-called *fast and frugal tree* (Martignon, et al 2003), that may be described in terms of three building blocks:

- (1) ordered search – features are used according to their “importance”,
- (2) fast stopping rule – the decision is made by a few steps as possible, and
- (3) one-reason decision-making – the single steps are based on one single criterion.

The tree asks only a few yes-or-no questions and has an exit at every level. If a patient has a certain anomaly in the electrocardiogram – the so-called ST segment is elevated – he is immediately classified as being at a high risk and is assigned to the coronary care unit. No other information is searched for. If that is not the case, a second question is asked. Is the patient’s primary complaint chest pain? If this is not the case, he is classified as low risk and put to a regular nursing bed. No further information is considered. If the answer is yes, a final question, namely whether any of four further main symptoms for cardiac infarction beyond chest pain is present or not, is asked to classify the patient and decide on the appropriate treatment.

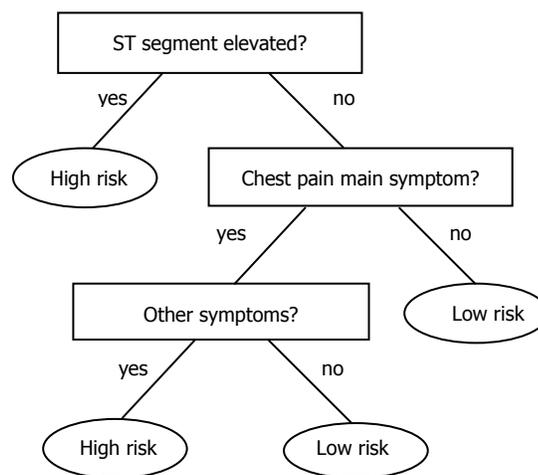


Figure 1. Fast and frugal tree to establish whether to assign a patient to the coronary care unit or to a nursing bed.

For a precise definition of the term “fast and frugal tree” and a mathematical treatment of such trees, see Martignon, et al. (2003) where the reader may find analogous trees used in the health context. We stress the point, that when physicians categorize patients based on one or two features only, they generally behave as natural Bayesians, that is Bayesians making use of natural frequencies (Gigerenzer & Hoffrage 1995), but when more features have to be used, they often switch to these fast and frugal trees.

Green & Mehr (1997) discuss the use of statistical techniques in medicine. They find that physicians prefer such simple routines over other, more complex statistical methods, mainly because of their transparency. This had been folklore knowledge for a long time, and the systematic analysis of these practices is fairly recent. What is astonishing is that the simple tree leads to more reliable decision processes than complex methods based on logistic regression (see also Martignon, et al 2008).

For medical decision routines, informed consent between doctors and patients requires a minimum of shared logical and stochastic literacy from both doctors and patients, in order to fully understand these routines and support decision-making. The medical realm is not the only one where these combinations of basic Bayesian reasoning and fast and frugal trees are systematically used. In more recent studies of the Group for Adaptive Behaviour and Cognition, financial decision making has been investigated in detail showing that shared decision making between advisors and clients often amounts to the use of these very simple trees (Monti, et al 2009). Thus, the dialogue between finance advisors and clients also requires a basic understanding of the probabilistic validity of investments’ features as well as the competency for ranking features, according to their validity.

An inference tool box for children

Preparing children for decision-making practices requires training them in making inferences, not just strictly logical but also, most importantly, probabilistic ones. They need to understand conditional probabilities for determining the validities of features and they need to be able to make comparisons between different validities of features for establishing rankings among features that produce fast and frugal trees to support decisions. These competencies are also at the core of risk assessment (Kurz-Milcke, et al 2008). Guided by research in the principles of human decision making, we devised the following sequence of simple tasks for fourth-graders:

- The inference “If your ST segment is elevated you require treatment in the coronary care unit” is a strict deterministic implication. Children have to grasp the essence of such implications. Thus, we begin with tasks fostering logical implications and deterministic inference.
- Children have to understand that the validity of a feature is established probabilistically. We thus extend strict deterministic inference to probabilistic inference and treat the task of assigning the probability of an item belonging to a given category, based on one feature (this probability is known as the validity or the diagnosticity of the feature).
- Decision trees such as the one described in Figure 1 are constructed by ordering features according to their validity. It is important to be able to compare validities of different features. Thus we treat proportion comparison in order to compare, for instance, validities of different features for categorization.
- As Bayesian reasoning is fundamental for computing validities of features, we let children work with urns and trees, which should prepare them to grasp the dependencies and perform the calculations.
- One application of probabilistic inference that children have to grasp as early as possible is reckoning with risk. We let children make use of comparing proportions for simple tasks, where they have to compare two situations, as to which is the riskier one.

Our first topic is that of logical implications. After letting children solve tasks that sharpen their understanding of the “modus tollens” rule for logical implication ^[1], we extend strict implications to *conditional* implications, that is, implications involving conditional probabilities. The implication of the kind

- “if P then Q”
is extended and transformed into a probabilistic statement of the type
- “the conditional probability of Q given P is high”.

These two statements are, of course, not equivalent even if the probability of Q given P were 1. We use concrete materials like cards and tinker cubes to support children in learning to distinguish between deterministic implication and probabilistic conditioning. We then let children apply probabilistic conditioning to categorization of units derived from the conjunction of features.

This practice is essential for decision-making. The next step is more mathematical and deals with the problem of comparing proportions. The ability to compare elementary proportions

may be acquired *before* learning fractions. (Fischbein, Pampu, & Minzat 1970). The comparison of proportions is essential for comparing feature validities and for assessing risks. In fact, we complete our experimental teaching with a game situation, in which the children have to compare such risks.

We tested our ideas by experimental teaching. After testing at other grades, we decided for grade 4 as optimal for our aims. Except for section 6, we had the same children participating in the whole series of our experiments, a total of 193 pupils in 6 classes; they come from schools in and near Ludwigsburg. The close links to such schools is a typical feature of that university where the first author works.

3. ELEMENTS OF LOGICAL THINKING: WASON CARDS

In this section we introduce a classical paradigm on how to empirically investigate people's understanding of logical "if-then" statements and report on studies implementing this paradigm and an extended version of it for the probabilistic implication in our experimental classes of grade 4.

The classical Wason experiment

A fundamental contribution to the understanding of how humans perform logical implications was provided by a series of experiments initiated by Wason in the late sixties (see Wason & Shapiro 1966). Wason tested subjects by showing them 4 cards on a table. The cards had letters on one side and numbers on the other. The subjects saw the following configuration:

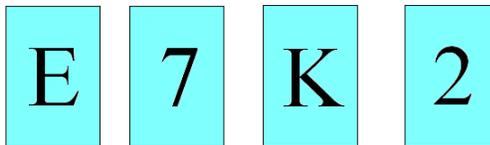


Figure 2. The original four cards of the Wason selection task.

Subjects were confronted with the following question: Which cards do you necessarily have to turn around in order to check, whether the following rule holds for the set of 4 cards?

“If one side of a card exhibits a vowel, its other side must exhibit an odd number”

The “E” card obviously might violate the rule, namely whenever an even number were to be found on the other side; therefore, its other side must be checked. Note that the rule

“vowel \rightarrow odd number”

is logically equivalent to

“no odd number \rightarrow no vowel” (which means “even number \rightarrow consonant”).^[2]

Therefore the “2” (and not the “7”) must be turned in order to check whether the other side eventually contains a vowel contrary to the rule. The “7” cannot violate the rule, since the rule implies no restriction for the back side of odd numbers.

It is reported that most subjects (> 85%) give a wrong answer. They either indicate cards E and 7, or card E alone. Only 13% of all subjects in the study notice that one has to turn around card 2 (in addition to the E card). Wason’s results seem to point to human flaws in logical reasoning.

Wason cards in contextual situations

This conclusion is contended by later experiments (Johnson-Laird, et al 1972), in which the abstract context of the binary features “vowels vs. consonants” and “even vs. odd numbers” is replaced by a context typical of social-contract situations. Let us look at a famous example of this type: each card represents an envelope to be sent by mail, with its destination written on one side and the value of its stamp on the other (Figure 3). Assume the rule is:

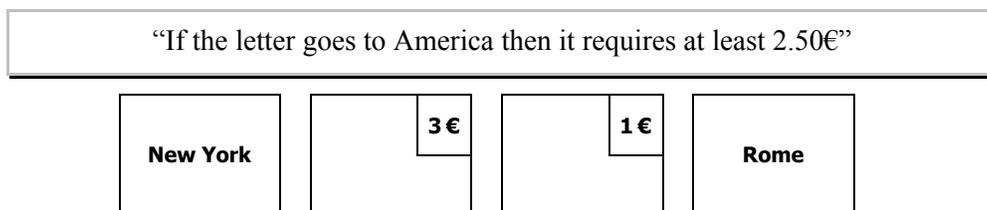


Figure 3. Four cards for an if-then rule from the “postal” contract.

Here the cards that have to be turned round are “New York” and “1€”, because a letter to New York requires at least 2.50€, and we have to make sure this one has such a stamp, and a letter with 1€ will never reach America. In tasks related to social contracts (the stamps required are established by postal decisions implicitly accepted by consumers), more than 85% are reported to solve the task correctly. A flurry of other experiments with rules from the social-contract context motivated the theory of the *cheating detection module*, on the one hand, and a historical approach to the evolutionary development of logic, on the other hand, as the normative system emerging from a “cognitive module” for cheating detection (e.g., Cosmides 1989).

According to Barkow, et al (1992), human flaws with implications in abstract settings are a mere sign of non-adaptation, like being unable to recognize colours under an unusual artificial light.

In our studies, we adapted the Wason task to a classroom setting and evaluated its educational potential; we transformed the task by modifying not just the context of the four cards, but also the number of cards used. By experimentally varying the number of cards (4 or 32) we modified the character of the original design. In our experimental classes, we let children sit around tables in groups of four and organize activities for 45 to 90 minutes (with a break of 10 minutes) with Wason cards. The four children at each table formed two teams.

The Wason cards we use include the original set with letters and numbers, sets with other symbols, colours, or short statements (like “If you help me in my homework” on the one and “you can borrow my bicycle” on the other side). In the first step, children worked with the coloured cards. In an explorative follow-up test, we studied whether the initial training with the coloured cards helped the children to solve tasks with Wason cards containing symbols or statements. In the typical situation of our studies, children get the opportunity to manipulate with the cards in their hands before the game starts. Thus, children could see both sides of the cards.

The design of our Wason experiments

- a. We used two pairs of colours that children usually see as pairs, namely “red vs. blue” and “black vs. white”. Our Wason cards were red or blue on one side and black or white on the other. Cards were thus red – white, red – black, blue – white, or blue – black, and children were aware of all the four possible combinations.
- b. The instructor wrote a “rule” on the blackboard:

“If one side of a card is red, then the other side must be white”

The six experimental classes were split in half according to the following two conditions:

Condition 1. Only four cards are placed on the tables in an array like in Figure 4. This is exactly the condition of the Wason task with the only difference that the cards have colours instead of letters or numbers.

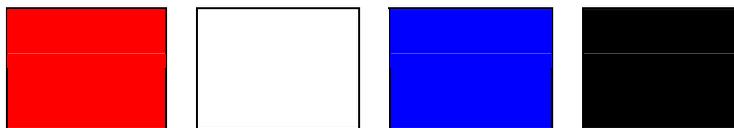


Figure 4. Four cards coloured either red or blue with reverse sides coloured either white or black.

Children have to determine which cards *need to* be turned around to check, whether the given rule holds. The instructor writes the following question on the blackboard, reads (loudly) and explains the game instructions:

“Which cards must be turned around in order to check that the rule holds?”

Condition 2. The instructor selects 32 cards out of a little red bag containing large amounts of cards and places them on each table; the selection is not done randomly as we have to ensure that the following cards are presented: 8 red – white, 8 red – black, 8 blue – white, 8 blue – black. The upper side is decided quasi-randomly with the constraint that each colour is on top at least four times. Again, the instructors write the question on the blackboard and explain the game instructions as in condition 1.

Alternating, each team must either point to a card to check – without actually turning it around – or state that no further cards have to be checked and so the game is over. The instructor acts as a “judge” and finishes the game if the current team has chosen a wrong card; in this case the other team wins. If the teams do not select any wrong card, that team wins that declares correctly that no further cards have to be checked. The team to start the game is decided by tossing a coin. When the game is over, the instructor flips the cards over and, if necessary, explains the “judgments” given during the game.

Let us briefly present the results. There were remarkable differences between the two experimental conditions.

- In condition 1, most teams (34 out of 36 teams) began by correctly flipping over a red card. At the next move, less than half (12) of all teams correctly suggested turning the black card around. The rest made an error: either they pointed at the white card or stated that the game was over.
- In condition 2, we observed that children were far more at ease. The game appeared to be more natural than under condition 1, and in only a few cases (10 out of 72 teams), a team eventually pointed to a “wrong” card, as one to be turned around, thus losing the game. The discussions at the tables after the game were partially reported by the assistants. Finally, all children realized that red – black cards were “forbidden”.

The long-term effect of learning was evaluated three weeks after the intervention with sets of cards with other contexts, e.g., the original Wason cards with letters on one side and numbers on the other. Again, children of the experimental condition 2 had significantly higher success than those under condition 1.



Figure 5. Playing with 32 Wason cards.

More than 80% of the children learning from condition 2 performed well three weeks after the intervention; this compares to roughly 40% of those learning from condition 1 who succeeded in that second test. We attribute the better performance to the different functions of the cards and the attractive game situation.

- Condition 1: Cards are seemingly experienced as abstract representatives of an *artificial* situation (four cards for the four cases ‘P and Q’, ‘P and not Q’, ‘not P and Q’, ‘not P and not Q’).
- Condition 2: The situation appears to be perceived as a real and engaging game with a lot of interactions between the children.

The observed differences in behaviour and success in the two conditions are indeed surprising. We only have tentative explanations. It is tempting to speculate that in condition 2 children are working with a sample and become *intuitive statisticians*. The experiments have been replicated in a looser form by teachers close to our group with basically the same results. It is important to note that in condition 2 children do not necessarily tend to uncover all “modus ponens” ^[3] cases first and then pass to the “modus tollens” cases: they act like detectives searching for a *thug* among their cards.

4. FROM CERTAINTY TO UNCERTAINTY

So far, we demonstrated how insight into the meaning of logical implications (“if-then” statements) might be fostered in young children. Changing the abstract context of the original Wason tasks into more intuitive contexts is one way, while increasing the number of cards might be another way to foster insight into an inferential situation that in the original setting is difficult even for adults. In this section, we illustrate our activities to let the children pass from strictly deterministic to probabilistic situations in which the concept of *proportions* becomes a crucial prerequisite for probabilistic reasoning.

Fischbein, among others, pointed out that young students tend to believe that “ambiguity and uncertainty are not acceptable in scientific reasoning”. In fact, this belief resembles that of adults in western countries during a large portion of our written history. Scientific argument, since Aristotle, was ruled by what is called classical logic until the Enlightenment when scientists of the calibre of Pascal or Laplace enhanced logical inference into a new instrument, capable of dealing with uncertainty, namely the probabilistic calculus (Daston 1988). One of the recommendations of the NCTM standards and the adaptations thereof in Germany is that children should be trained in reasoning and argumentation from an early age, not just in mathematical contexts. Reasoning comprises not just practicing deduction on absolute certainties but also drawing “inductive” conclusions from uncertain phenomena.

In this paper we want to stress the importance of logical argument and the importance of a conscious transition from certainty to uncertainty, which will characterize not just the study of stochastics but a general form of reasoning. Following a “historic” trajectory we let children experience the transition – from logic to probability – in our exploratory studies.

A looser form of implication and Wason cards

In our experimental classes of grade 4, we discussed with the children that

- “if-then sentences” have limitations,
- that such statements may fail to provide adequate descriptions of real situations, and that
- a “looser” form of conditioning describes real-life situations better.

In two subsequent 2 hour blocks, we let “our” children experience the transition from Wason cards to tinker cubes for representing individuals. Children were again organized in teams of two and sat at tables. We used the following introductory statement:

“If you are a boy, then you like computer games”.

We provided plain cards to all children, so that they could produce the Wason cards for this rule. The aim was to examine Wason cards for two features like the following:



Figure 6. Wason cards with if-then rules describing children’s preferences in subjects.

Questions written on the blackboard were designed to motivate the use of cards for representing individuals. The questions were:

1. Is it true *in this class* that if you are a boy you like computer games?
2. Is it true *in this class* that if you are a girl you do not like computer games?

Children were asked to state whether they liked computer games and colour their card accordingly. Sets of cards to be coloured were placed on the individual desks of each child. They had to colour their own cards according to their values so that each child was represented by its card. We used the coding scheme in Table 1.

Cue	Code	Cue	Code
boy	blue	Likes computer games	white
girl	red	Does not like computer games	black

Table 1. Coding schemes for colours.

A boy who likes computer games is represented by a *blue – white* card, while a *red – black* card represents a girl who does not like computer games. Once the class was represented by a set of cards on the desk of each child it became easy to count and calculate statistics. Large two by two tables were drawn on sheets of paper which were also placed on children’s desks. Children had to distribute their cards in the four cells of their two by two tables, according to the encoding in each card. Teachers eventually passed by and checked for the number of cards in each cell. This practice was successful in the sense that children were good at “constructing” their class and classifying it with the help of the 2×2 table on the large sheet.

Conditional probabilities and tinker cubes

Wason cards are symmetric, in the sense that none of the sides is privileged, and they can only encode two binary ^[4] features, of which one is visible and the other not, when they lie on a

table. They provide a mystery situation that converts children into “detectives”. It is interesting to swap from Wason cards to cubes and towers assembled with cubes, because these characteristics – symmetry and mystery – are replaced by new characteristics that introduce a new flavour to the activities. *Tinker cubes* is the name we gave to small cubes made of plastic in different colours that can be assembled to form tinker towers that encode feature information. For instance, a “tinker tower” consisting of one red and one black cube might represent a girl (red) who does not like computer games (black).

Although these coloured plastic cubes have been around for decades in most countries of the world, we decided to give a special name to these cubes when used with this particular context, namely as units that can encode information and can be assembled to form individual profiles. Of course, tinker cubes can not only display combinations of two dichotomous variables but can also easily be used for displaying all sorts of simple frequency-distributions. In Figure 7, tinker towers are used to encode data on fruit preferences of children, thus forming a tangible histogram. Here the blue tower is composed of cubes each of which stands for one child preferring banana in two parallel fourth classes.



Figure 7. Tinker towers for fruit preferences.

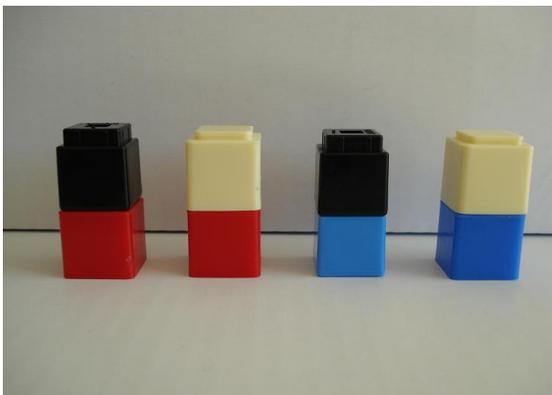


Figure 7a. Tinker towers representing conjunctions of 2 ...

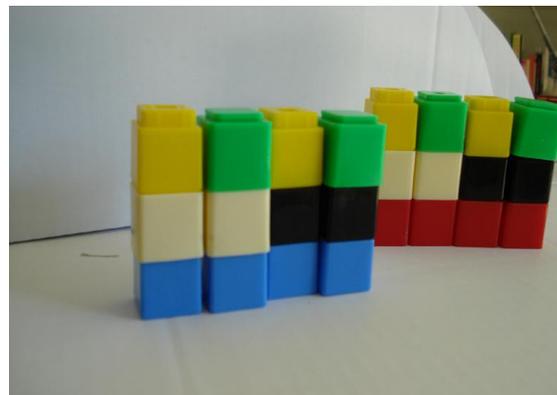


Figure 7b. ... of 3 features.

Tinker towers, as already mentioned, produce a situational change with respect to Wason cards, for several reasons, one being that all feature colours are visible at the same time. Furthermore, tinker towers can be seen as ordered “vectors” of values, represented by colours. In fact, tinker cubes have a bottom and a top. Thus, they may be used for encoding the values of binary features in a given order. We taught the children in our experimental classes en-actively to re-construct statistical situations with the help of tinker towers [5].

One activity (step 1), already mentioned above, was the construction of children’s “fruit preferences”. In another activity (step 2) 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*, in analogy to the coding scheme with the Wason cards. In this situation, it is easy to add further features like: *having or not having mathematics as one’s favourite subject*. The pair *yellow-green* was used to encode this new feature (Figure 7b).

The next step (step 2) was to perform quantified categorizations and introduce the concept of *proportions*, which lays the first foundation for the concept of *conditional probability*. The instructor asked, for instance, what proportion of girls likes computer games. Such statements will later be transformed to probabilistic questions, see Table 2.

Proportions	Probability
What proportion of girls likes computer games?	Given a pupil is a girl, what is the probability that she likes computer games?
What proportion of those who like computer games are boys?	Given a pupil likes computer games, what is the probability that this is a boy?

Table 2. Translation of statements in proportions to statements in probability.

Evaluation of the activities

At the end, a written test was performed. Children were given sheets with data on two fourth-grade classes in the United States. Here they had to deal with two features or cues:

1. *To like or not like* computer games
2. *To like or not like* mathematics as one’s favourite subject

Constructing, sorting, counting and assessing proportions of tinker towers in *nested* sets of feature conjunctions is an activity that children perform with ease and pleasure and marks a first step towards a fruitful combination of descriptive statistics on one hand and “natural”

Bayesian reasoning on the other. Most children ($\approx 70\%$) in each of the six classes were successful in dealing with all kinds of conditional assessments and could answer questions like “Do boys who like mathematics necessarily like computer games?”, or “Are those who do not like computer games and yet prefer mathematics over other subjects more likely to be girls?” The term “proportion” was also consistently used, in questions like “what proportion of children with short hair are boys?” Children had to find out, how valid a cue like “long hair” is, as a feature of the category “girl”, or how valid the feature “wears trousers” is for the category “boy”. Children learned the concept of validity of a feature and learned to express it as a conditional proportion.

5. COMPARING PROPORTIONS

Finding the validity of *one* feature only is not enough for good decision making. In most relevant cases, decisions are not based on one feature but on several. As an example, recall that breast cancer is not diagnosed solely on the basis of the result of a mammography, but on a whole chain of results, like mammography, ultrasound, and tumour markers. Green & Mehr (1997) report that humans seem to rank features according to their validity and tend to consult them, one by one, until they reach a decision. Thus, being able to compare proportions is crucial for decision making; not just because it becomes the instrument for ranking features but also it is an eminent competency per se. Interventions on proportional thinking establish a necessary component in a tool box of stochastic activities for children.

Playful activities for comparing proportions

We deliberately chose to perform these interventions on the comparison of proportions *after* those on conditional probabilities. In our classes, we performed experimental teaching in blocks of 6 hours, in proportional thinking, and found that fourth-graders can assimilate training in proportions quite well. After several pilot experiments in classes of grades 3 to 5, we have come to the conclusion that fourth grade is perfectly suitable for introducing “proportional” activities in a playful way.



Figure 8a. Green and red gum bears.



Figure 8b. Other proportions.

If urns are filled with sweets in the composition of either the left or the right picture in Figure 8, one may ask:

Assume that you have to pick a bear by drawing blindly from one of two urns in Figure 8. Which urn is “favourable,” if you want a red bear?



Figure 9a. Red and green bears.



Figure 9b. Different proportion of red and green.

This is an easy task, children say. The second urn is better. But things can get cumbersome if we have to compare the choices in Figure 9. Training children in this kind of tasks is equivalent to the problem tested by PISA 2003:

“From which urn should you draw ‘blindly’, if you want a white marble, given the situation in Figure 10?”

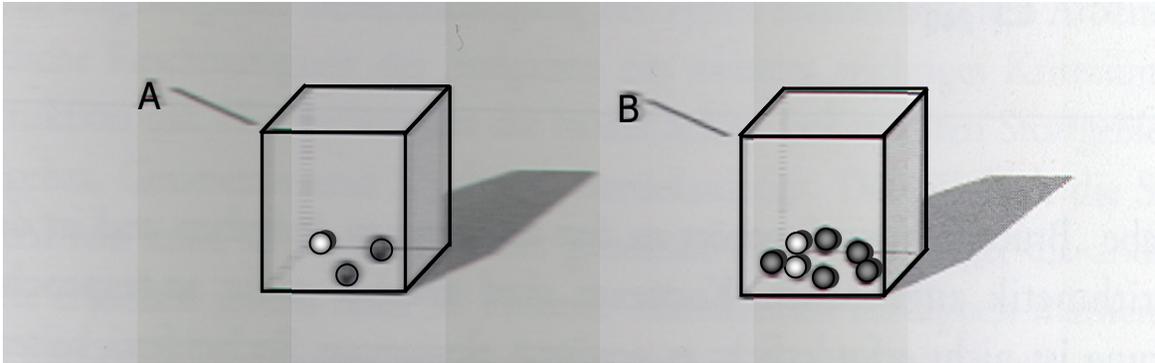


Figure 10. Urn 1: 1 white and 2 black marbles – Urn 2: 2 white and 5 black marbles.

We recall that the performance of German students (15 years old) was particularly poor for this task: only 27% solved it correctly (Pisa Consortium Deutschland 2004). One should remark how relevant this type of procedural competency is, namely to establish whether 1 out of 3 is larger than 2 out of 7. This competency enhances our perceptual capacity.

Enhancing the comparison of proportions

The investigation of children's ability to deal with proportions is quite an old field of research. Fischbein, Pampu, & Minzat (1970) devised beautiful experiments to evaluate children's success that have inspired the design for our own experimental units (see also Spinillo & Bryant 1991). Assume, for instance, that we compare the ratio of orange to blue squares in the setting of Figure 11.

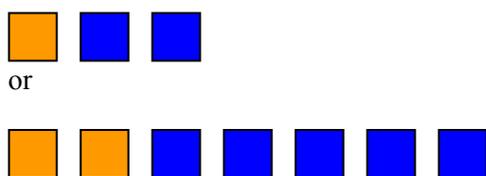


Figure 11. Urn composition, for the urns of Figure 10, now in orange and blue cubes.

The question remains the same as is the PISA setting:

Which array is “favourable” if you want to pick an orange cube by drawing blindly from one of them?

The school of Fischbein suggested working with “similarities”, that is by adequately translating the given proportions to other, *equivalent* proportions that can be more easily compared.

There is controversial evidence in the literature about the success of proposed strategies to transform proportions preserving their “value”. Piaget (1951) pointed out that young children are prone to the additive fallacy according to which they tend to think that the proportion “1 to 2” does not change if we add 1 to each side, obtaining the ratio “2 to 3”. In their work with indigenous people in the Amazon region, Dehaene, et al (2006) show that such a constancy of proportions is an archetypical kind of thinking, which may be noticed in people without schooling at all [6]. They used *geometric* representations in their experiments as in Figure 12.

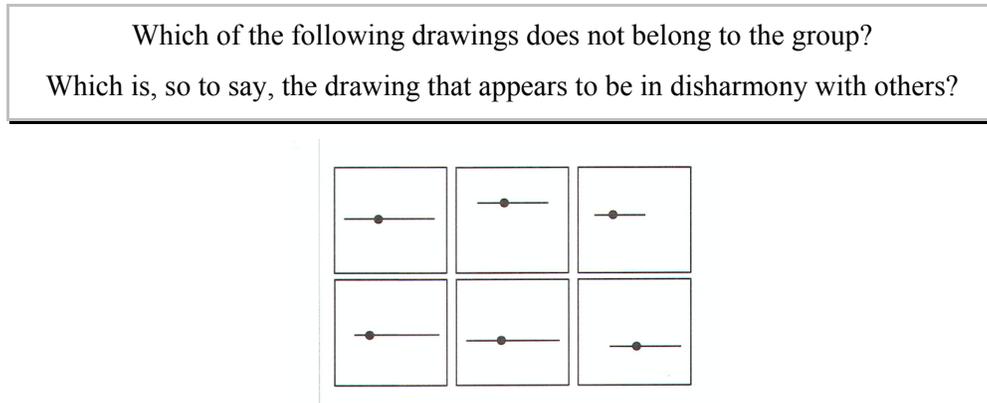


Figure 12. Geometric proportions from a task devised by Dehaene, et al. (2006).

The additive bias of Piaget – if it existed – may also be met by special experiments, which were designed by Stern and Koerber. They base their research on a remarkable capacity to detect changes in flavour. In a mixture of e. g. Coca Cola and lemonade (in Germany known as “Spezi”), if an equal amount of each ingredient is added, this changes the flavour and children are very good in detecting such changes. If the activities of mixing and adding the same amount are graphically represented, the children are supported to recognize and acknowledge that an equal amount added to both components changes the *flavour* and therefore “should” also change the *proportions* in the geometric representation (see Stern, et al. 2002, Koerber 2003).

We have therefore some support for our choice of a markedly geometric approach to the comparison of proportions in the classroom. We, or in some cases the teachers, instructed the children to construct towers “similar” to given towers, like that in Figure 13a.

Children had to add orange and blue cubes to this tower keeping the geometric constancy in the proportion of the orange part with respect to the blue part. The term “geometric constancy” was not pronounced: children had to judge “similarity” in the proportions of orange and blue in the towers by the visual impression of the proportion. The towers in Figure 13b, for instance, preserve this proportion.

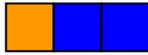


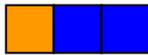
Figure 13a. Basic tower.



Figure 13b. Towers equivalent to the basic tower.

Children constructed many “similar” towers and were good at checking for “similarity”. We noticed that most children enjoy forming ever larger similar towers preserving the ratio “1 to 2” (corresponding to the statement “1 out of 3” cubes is “orange”). See Figure 14 for the kind of operations that preserve the initial proportion and finally lead to a situation, in which the given two proportions of 1:2 and 2:5 are easy to compare directly.

If we know that



1:2

and



2:4

are the “same”, all we need to find out is whether



2:4

contains proportionally less orange cubes than



2:5

to conclude that 1: 2 is better (for orange) than 2:5.

Figure 14. Steps to construct equivalent towers in order to compare the proportions of 1:2 and 2:5.

Evaluation of the class activities

In our experimental classes we worked consistently along these lines letting children en-actively compare proportions. The lasting effects of teaching were evaluated four months later by written tasks, where the children could use their tinker cubes and towers as a support for solving the tasks. Tinker cubes were placed in baskets close to children’s desks and they could pick as many as they wanted to work on their desks.

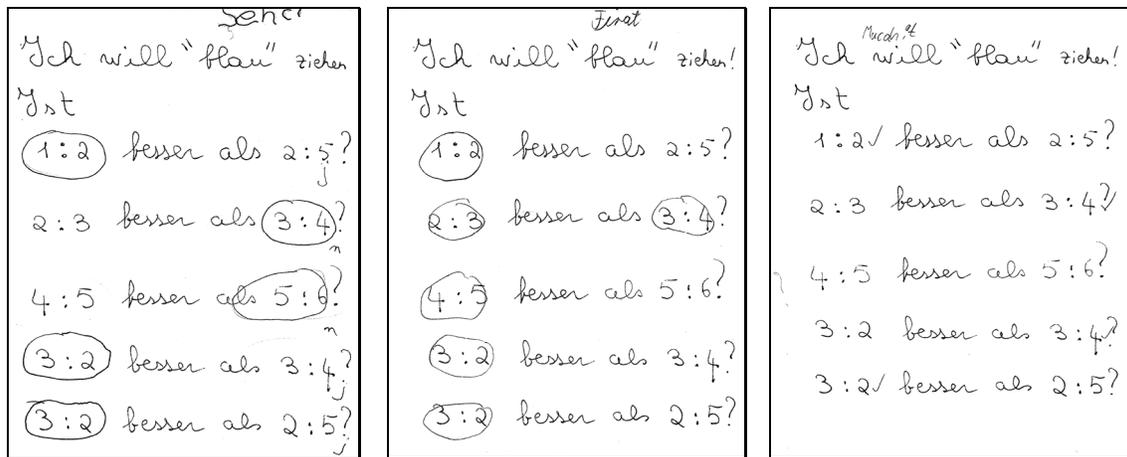


Figure 15. Three tests of Sehci, Firat, and Matchet illustrating their work.

In these tasks, children had to compare proportions and indicate, which of two proportions contained “relatively more” blue cubes, where the number of blue cubes was always at the left side of the proportion. The amount of blue cubes corresponded to the quantity on the left hand side of each ratio. The success of our units is not only represented by children’s ability to solve these tasks (more than 60% in each class had more than 3 out of 5 items correct), but also by their enthusiasm in solving and checking their results by constructing tinker towers and comparing them.

6. FROM PROPORTIONAL TO PROBABILISTIC THINKING

In modern life, we have to estimate risks based on features or characteristics, and these features may not be completely trustworthy, like medical tests, or expert ratings. A medical test may sometimes fail to detect a disease, or it may err by being positive even in cases where the patient does not suffer from the disease (“false positive”).

What are the chances of being HIV positive when one has a positive Eliza Test?

The main question is: how diagnostic is such a test – or, less technically, how reliable is a diagnosis based on this medical test?

The importance of representations for probabilistic thinking and Bayesian reasoning

Here again one has to acknowledge that the formats used for dealing with such matters can be cumbersome and difficult to grasp, as Gigerenzer (2002) has emphatically claimed, or

“natural” and transparent. Human cognition is better tuned to working with so-called “natural frequencies” than with modern probabilities. Proportions like “8 out of 100” are better understood than expressions like 8%, quotients like $8/100$, or decimals such as 0.08. But, what is more important, if the information about the sensitivity and the specificity of a test (like Eliza) as well as the base rate of a disease (like HIV) are presented in natural proportions, then humans have no difficulty in estimating the risk of actually having the disease.

This, apparently, has to do with the greater ease in processing information when *adaptive* inner visual representations (like imagining 8 persons in a group of 100) are triggered. It even turns out that probabilistic information represented by icons so as to simulate natural frequencies (for individuals in populations and subpopulations) is processed in a more effective way than probabilistic information provided by means of Venn diagrams. Without going through the whole discussion of this phenomenon, we refer to the recent work of a cognitive psychologist (Brase 2008), who presented the following three stimuli as in Figure 16.

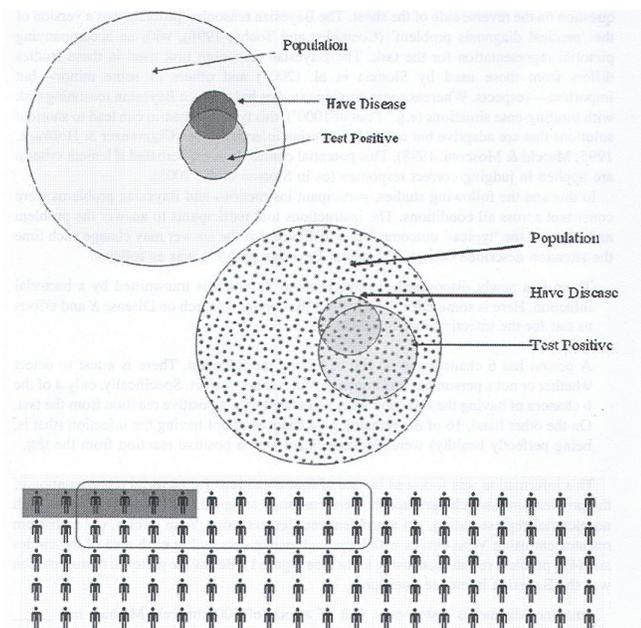


Figure 16. Three stimuli presented by Brase (2008).

Even from the perspective of visual perception, the stimulus provided by icons (of people), with shaded proportions of selected subjects, is the most effective, i. e. it is the one most easily processed. Based on this type of research, we work persistently with enactive representations, like tinker cubes and tinker towers that have this same discrete character, where tangible units encode not just individuals but their features.

Urns and trees proportional reasoning

Such an early preparation in fourth grade should provide anchoring for later education in formal conditional probabilities and Bayes' theorem. Kurz-Milcke designed and performed an intervention block with urns and tinker cubes in the fourth grade: These grade 4 children constructed urns containing binary tinker towers encoding binary features (like going to movie A or movie B, or going to sports arena A or B). Children also constructed *trees* on the floor, as the one depicted in Figure 17b, where the first level represented the two urns for one of the two variables treated, and the second level partitioned each urn further into two urns corresponding to the second binary variable. For detailed results of that experiment see Kurz-Milcke & Martignon (2006); note that this experiment is the only one that involved a class different from our six experimental classes.



Figure 17a. The urn of "our" class.^[7]

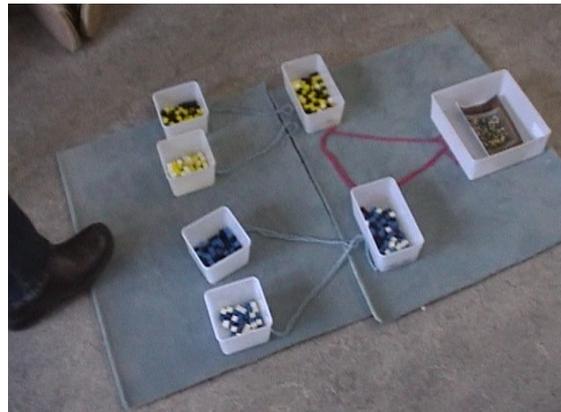


Figure 17b. Trees for tinker towers encoding 2 binary variables.

Proportions of tinker towers may be used as enactive representations of proportions of subpopulations of individuals – with certain conjunctions of features – out of larger populations. These proportions are a first step towards probabilistic thinking.

With drawing balls from urns, usually two functions are associated:

- a. Urns as static containers: urns are mere containers of balls that can be counted.
- b. Urns as a dynamic source of changing results: urns are "shaken" to mix up the balls for random drawing.

In our typical interventions (Kurz-Milcke & Martignon, 2006; Martignon & Kurz-Milcke 2006) plastic urns were filled with tinker towers. The urn was "shaken" by one child while another child drew "blindly" from it; shaking urns and drawing blindly from them adds a probabilistic aspect to tinker towers. Thus, urns were used consistently with their two functions: as containers and as mixers of their content.

7. FROM COMPARING PROPORTIONS TO COMPARING RISKS

In this last section we suggest how to introduce young children into the issue of understanding risks, which usually are communicated in terms of probabilities or percentages. In doing so, we again rely on children's understanding of proportions. Although it is not possible to introduce the modern concept of probability (or conditional probability) in grade 4, we argue that by establishing a link from probabilities to proportions, understanding of risks can be fostered. Reckoning with risk has become all the more subtle for humans in modern times since most risks are communicated by means of mathematical formats that have to be learned in school. Both medical and finance decision making is based on statistical information.

Thus, training young students in the perception of risk has become fundamental in modern society. Students have to acquire tools to model the risk of having a disease like AIDS, or of being pregnant when the pregnancy test is positive, of losing money with an investment, etc. Expressed in mathematical terms, a risky event is one associated with a strictly positive probability of a loss of resources like health, time, food, or money. These resources are usually modelled in terms of "utilities".

Risk analysis in the game Ludo

The framework for understanding risk can be taught in secondary school, but the basic intuitions can – and should be – fostered in primary school. Stochastic activities in primary school cannot be limited to urns, albeit these are the basis of a good stochastic training. Young students have to learn to reflect on resources and loss of resources. This is what behavioural scientists and decision theorists would recommend. Fortunately, there are many popular board games for children involving possible losses and gains of resources. These games can be used to illustrate risky behaviour. For instance, German children begin playing Ludo (in German "Mensch ärgere Dich nicht") when they are about five years old. We give a short account of the game's instructions:

Ludo: 2–4 players take it in turn to throw a single die. A player must first throw a six to be able to move a token from the starting area onto the starting square. In each subsequent turn the player moves a piece forward 1 to 6 squares as indicated by the die.

When the die shows a 6, the player may bring a new piece onto the starting square, or may choose to move a piece already in play. Any throw of a six results in an extra turn.

If a player cannot make a valid move, the die is passed on to the next player.

Only 21 of the 193 children (less than 11%) wrote that they would choose option 1 – to move the left green token – and gave a conclusive explanation. It seems that children and adults as well tend to move forward that token that is the furthest ahead. In the game situation presented, this is the right green token as one can see from Figure 18 and the arrow indicating that the flow of the circuit is done clockwise. In this habit they ignore the higher risk of losing a token at the next move of the adversary that is related to their choice. Risk proneness and risk aversion may be illustrated to children in this simple set up. Of course, there could also be something in background that affected the decision, which might be summarized by the saying “Nothing ventured, nothing gained” (in German “Nur wer wagt, gewinnt.”)

Relative and absolute risks

In the situation described above, the risk of losing one token can have a probability of “one out of three” if one moves further with the right token (Figure 18) or “one out of six” if one moves with the left token. How might one compare the risk in the two scenarios?

Is it correct to say, that
 “the expectation of Green keeping a token has been increased by 100%”,
 when actually
 “the risk of losing one token has been reduced from “2 out of 6” to “1 out of 6”?

This question may seem awkward or even trivial and yet it is central to the understanding of risks in the way they are communicated in modern media. Both in the context of finance and medicine “relative” risks are often presented instead of “absolute” risks, thus blurring the understanding of their real impact. In the above example, the problems of absolute and relative comparisons are multiplied as the statements switches to the complementary event of a token being captured, namely that the player prevents the token being captured. Table 3 gives an overview on the changes of the risk when options 1 and 2 are compared.

Risk that one token is captured in the next move	Compared to the other option			Compared to the other option	
	Change of risk absolutely	Change of risk relatively	Probability for keeping the tokens	absolute	relative
Option 1	17%	-17%	83%	+17%	20%
Option 2	34%	+17%	66%	-17%	25%

Table 3. Absolute and relative risks of several options in the game situation.

As another example, the following type of information is 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 the drug, which the commercial recommends? (from Doctor's Guide n. d.):

Heart-Attack Risk Reduced by almost Half with Aggrastat

This is no singular case; one can find plenty of similar commercials communicating risk reduction when using this or that drug. We mention two further examples. One is related to the famous Lipitor, a medicine for the prevention of cardiac infarction – again there is the promise of risk reduction communicating both absolute and relative risks, yet absolute risks in large characters, while information on relative risks is written in tiny characters in a footnote.

An advertising campaign by the pharmaceutical company Pfizer focused on the core argument of reducing the risk for heart disease by taking Lipitor regularly:

The dramatic 36% figure has an asterisk, which was *added* in the advertisement in much smaller letters. It says: "That means in a large clinical study, 3% of patients taking a sugar pill or placebo had a heart attack compared to 2% of patients taking Lipitor." Meanwhile the campaign has been withdrawn – that is why we cannot present the original picture to give the reader an authentic impression how the various risks were used (see, however, [Mouseprint](#) n. d.)

"In patients with multiple risk factors for heart disease,
Lipitor
 reduces the risk of
heart attack
 by **36%***
 If you have risk factors such as family history, high blood pressure, age, low HDL ('good' cholesterol) or smoking."

Figure 19: Advertisement for Lipitor
 – see e. g. [Entrepreneur](#) (n. d.).

The other example is just another instance of the same phenomenon: Women are constantly prompted to perform regular screening by mammography to enhance their life expectancy by 25%. How should this enhancement be interpreted? The 25% enhancement of life expectancy may be explained as follows: 4 out of 1000 women have breast cancer, and if these 1000 women perform screening regularly then one will be saved. It is crucial to become aware that the risk is reduced from 4 out of 1000 to 3 out of 1000, which amounts to an risk reduction of *absolutely* 1 out of 1000, i. e. 0.1%, but *relatively* the reduction amounts to 1 in 4, i. e. 25%. Gigerenzer (2002) has looked at this and many other examples and the dramatic consequences of the confusion between “relative” and “absolute” risk.

How can schools prepare for decision-making and shared informed consent? How can schools prepare future grown-ups to be alert, when the media promote products or practices by communicating relative risks instead of absolute risks? The answer is clear if complex: our school curricula have to provide basic tools for dealing with an uncertain world.

8. CONCLUSIONS

Germany's educational system is beginning to devote more time and attention to teaching data analysis and probability. In fact, competencies in data analysis and probability are now a mandatory component of school curricula from elementary school to grade 12. The University of Education at Ludwigsburg, for instance, provides mandatory courses in stochastics and its didactical aspects to all future mathematics teachers of the first school years.

But what exactly should be taught, say, in elementary school, in order to stimulate probabilistic reasoning?

The studies we present here investigate an approach to build a solid basis for the later acquisition of the calculus of probability. We propose the consistent use of materials like cards, tinker cubes, and tinker towers for training in basic inference and proportional thinking. Our first analyses indicate that proportional and probabilistic reasoning based on hands-on activities with these tools constitute a successful step towards probabilistic comparisons for decision-making and reckoning with risk. Our results encourage further experiments in that direction. It is promising to see that some of our collaborating teachers have quickly integrated our practices into their standard repertoire and develop their own units, with their own creative approaches encouraging children to work with tinker cubes and tinker towers.

REFERENCES

- Amit, M. & Jan, I. (2007). *A learning environment that promotes concept construction: The case of probability*. Paper presented at the annual conference of the American Educational Research Association, Chicago, IL.
- Barkow, J. H., Cosmides, L., & Tooby, J. (Eds) (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.
- Brase, G (2008). Pictorial representations in statistical reasoning. *Journal of applied cognitive psychology*, 3, 1002–1016.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31, 187–276.
- Daston, L. (1988). *Classical Probability in the Enlightenment*. Princeton; NJ: Princeton University Press.
- Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core knowledge of Geometry in an Amazonian indigene group. *Science* 311 (5759), 381–384.
- Doctor's Guide (n. d.) Retrieved from <http://www.pslgroup.com/dg/7DEB6.htm>, March, 25, 2009.

- Entrepreneur (n. d.). *Latest heart news: the good, the bad, and the not so bad*. Retrieved from <http://www.entrepreneur.com/tradejournals/article/173486299.html>, July 7, 2009.
- Finkel, A. (2008). Still a Task for Sisyphus. *Annals of the New York Academy of Science*, 1128 (1), 121–137.
- Fischbein, E., Pampu, I. and Minzat, I. (1970). Comparison of ratios and the chance concept in children. *Child Development*, 41, 377–389.
- Fischer, J. E., Steiner, F., Zucol, F., Berger, C., Martignon, L., Bossart, W., Altwegg, M., & Nadal, D. (2002). Use of Simple Heuristics to Target Macrolide Prescription in Children with Community-Acquired Pneumonia. *Arch Pediatr. Adolesc. Med.*, 156, 1005–1008.
- Gigerenzer, G. (1998). Ecological intelligence: An adaptation for frequencies. In D. D. Cummins & C. Allen (Eds.), *The evolution of mind* (pp. 9–29). Oxford: Oxford University Press.
- Gigerenzer, G. (2002). *Reckoning with risk: learning to live with uncertainty*. London: Penguin Books.
- Gigerenzer, G. & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1), 107–144.
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L., & Woloshin, S. (2009). Helping Doctors and Patients make sense of Health Statistics. *Psychological Science in the Public Interest*. 8-N. 2, 53–167.
- Gigerenzer, G. & Hoffrage, U. (1995) How to improve Bayesian reasoning without instruction: Frequency Formats. *Psychological Review*, 102 (4) 684–704.
- Green, L. & Mehr, D. R. (1997), What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice*, 45(3), 219–226.
- Hirsh-Pasek, K. & Golinkoff, R. (2003). *Einstein never used flash cards: How children really learn and why they need to play more and memorize less*. Emmaus: Rodale Press.
- Johnson-Laird, P.N., Legrenzi, P., & Legrenzi., M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, 63, 395–400.
- Koerber, S. (2003). Visualisierung als Werkzeug im Mathematik-Unterricht: Der Einfluss externer Repräsentationsformen auf proportionales Denken im Grundschulalter. Hamburg: Dr. Kovač.
- Kurz-Milcke, E., Gigerenzer, G., & Martignon, L. (2008). Transparency in Risk Communication. *Annals of the New York Academy of Science*, 1128 (1), 18–28.
- Kurz-Milcke, E. & Martignon, L. (2006) Lebendige Urnen und ereignisreiche Bäume: Überlegungen und Versuche zu einer Didaktik der Stochastik in der Grundschule. In J. Meyer (Hrsg.) *Anregungen zum Stochastikunterricht, Bd. 3.* (pp. 181-203). Hildesheim: Franzbecker.
- Martignon, L., Katsikopoulos, K., & Woike, J. (2008). Categorization with simple heuristics: Fast and frugal trees. *Journal of Mathematical Psychology*, 52, 352–361.
- Martignon, L., Katsikopoulos, K., & Woike, J. (2010). Naive, fast and frugal decision trees. In P. Todd, G. Gigerenzer, & the ABC Group. *Ecological rationality: Intelligence in the world*. Oxford: Oxford University Press.
- Martignon, L. & Krauss, S. (2007). Gezinkte Würfel, Magnetplättchen und Tinker Cubes für eine Stochastik zum

- Anfassen in der Grundschule. *Stochastik in der Schule*, 27, 21–30.
- Martignon, L. & Kurz-Milcke, E. (2006). Educating children in stochastic modeling: Games with stochastic urns and coloured tinker cubes. In A. Rossmann & B. Chance (Eds.), *Working cooperatively in statistics education. ICOTS 7 – Salvador de Bahia*. Retrieved from <http://www.stat.auckland.ac.nz/~iase/publications/17/C443.pdf>, June 22, 2009; also available on CD.
- Martignon, L., Vitouch, O., Takezawa, M., & Forster, M. (2003). Naïve and yet enlightened: From natural frequencies to fast and frugal decision trees, In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological Perspectives on Reasoning, Judgment, and Decision Making* (pp. 189–211), Chichester: John Wiley and Sons.
- Monti, M., Martignon, L., Gigerenzer, G., & Berg, N. (2009). Simple strategies for investment decisions. Submitted to *Journal of Economic Psychology*.
- Mouseprint (n. d.). *Consumer World*. Retrieved from http://images.google.de/imgres?imgurl=http://www.mouseprint.org/wp-content/avert/lipitor2.jpg&imgrefurl=http://www.mouseprint.org/2008/02/04/lipitor-reduces-bad-cholesterol-but/&usq=__guEHVeGfsG_bqbpBqVA8bQ5EGOY=&h=399&w=500&sz=60&hl=de&start=5&um=1&tbnid=D6d, July 15, 2009.
- Neubert, B. (2007). Kompetenzen von Grundschulern bei der Bearbeitung von Aufgaben zur Wahrscheinlichkeitsrechnung, *Stochastik in der Grundschule*, 27, 28–33.
- Peters, E. (2008) Numeracy and the Perception and Communication of Risk. *Annals of the New York Academy of Science*, 1128, 234–267.
- Piaget, J. & Inhelder, B. (1951). *La genèse de l'idée d'hazard chez l'enfant*. Paris: Presses Universitaires de France.
- Pisa Consortium Deutschland (2004). *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland-Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Spinillo, A. G. & Bryant, P. (1991). Children's Proportional judgments: The importance of "half". *Child Development*, 62, 427–440.
- Stern, E., Hardy, I., & Koerber, S. (2002). Die Nutzung graphisch-visueller Repräsentationsformen im Sachunterricht. In: K. Spreckelsen, A. Hartinger & K. Möller (Eds.), *Ansätze und Methoden empirischer Forschung zum Sachunterricht* (119–131). Bad Heilbrunn: Klinkhardt.
- Wason, P. C. & Shapiro, D. (1966). Natural and contrived experience in a reasoning problem. In Foss, B. M. (Ed.), *New horizons in psychology*. Harmondsworth: Penguin.
- Zhu, Y. & Gigerenzer, G. (2006). Children can solve Bayesian problems: the role of representation in mental computation. *Cognition* 98, 287–308.

APPENDIX – CUES, RANKING OF CUES, AND VALIDITY OF CUES

While statisticians think in *variables* as measurable characteristics of statistical units, psychologists would talk about *features or cues*. Cues can be symptoms, characteristics, or test results. Quite often these features are binary (positive or negative).

Let us give an example from medical context. A woman may have breast cancer D or not \bar{D} . A cue for breast cancer is e. g. anomalies in the mammography that result in a so called “positive mammography”. One then expects:

$$p(D | \text{mammography anomalies}) > p(D).$$

Breast cancer is the *target variable*, which is not directly amenable to “measurement”. There are other variables, or features, which are more or less suitable to diagnose the disease. Mammography is but one, ultra-sound another, several tumour markers establish further “cues”.

The results of indirect measurements are used for classifying a patient to have or not have the disease based on the test results. The positive value (+) is used for diagnosing the disease.

- *Valid cue* means that the presence of this cue increases the conditional probability of the disease to a high value (or at least increases the probability to a much higher value than without this cue).
- *Invalid cue* means that the presence of this cue decreases the conditional probability, i. e. it increases the complement’s probability.
- *Invalid cue* sometimes is also used to denote the circumstance that the conditional probability of the disease is only slightly increased; or, that it is increased but still is a comparably small probability.

If there are indicator variables J_1 with $+_1$ and $-_1$, J_2 with $+_2$ and $-_2$ etc., with the following ranking:

$$p(D | +_1) > p(D | +_2) > p(D | +_3) \dots$$

then we say that there is a ranking with respect to the validity of the cues in the way that

$$J_1 > J_2 > J_3 \dots$$

Usually *prior probabilities* are *assumed*, which are appropriate for the whole population or for a subgroup to which the examined person belongs (e. g. a risk group because of inheritance factors for the disease).

The reversed probabilities are in general more or less known by practice. The pharmaceutical industry is used to assessing sensitivities and specificities of tests.

For appropriate applications in decision-making, the probabilities of the disease have also roughly to be known. The calculations can be done by the Bayesian theorem (here only for the k th cue)

$$p(D|+_k) = \frac{p(D) \times p(+_k|D)}{p(D) \times p(+_k|D) + p(\bar{D}) \times p(+_k|\bar{D})}$$

from the prior probabilities	the sensitivities	and the specificities
$p(D)$	$p(+_k D) \dots$	$p(-_k \bar{D}) \dots$
of the disease	of the indicator variables.	

Note that the validity of a cue *cannot be computed without knowing the prior probabilities* of the values of the target variable. Thus, a ranking of validities of cues cannot be established without knowledge of prior probabilities, sensitivities and specificities of all cues. The prior probability, or *incidence* of a disease, is usually estimated by health centres. The incidence of breast cancer in women older than 20 years in the US has remained fairly stable during the last decade. The ABC Group postulates that validities are assessed using natural frequencies rather than Bayes' formula. Humans are "natural" Bayesians, when establishing the validity of *one* cue, and use fast and frugal trees, when they have to make a classification based on several cues. Regardless of the format used for Bayes' theorem, schools have to start preparing children sufficiently early.

Author : Laura Martignon
 E- Mail : martignon@ph-ludwigsburg.de
 Address : University of Education, Ludwigsburg, Germany

Author : Stefan Krauss
 E-Mail : stefan1.krauss@mathematik.uni-regensburg.de
 Address : University of Regensburg, Germany

NOTES

-
- [1] The modus tollens rule reads as follows: $(P \Rightarrow Q)$ and *non* Q *implies non* P
- [2] This is called the law of contraposition, which amounts to modus ponens.
- [3] The modus ponens rule reads as follows: $(P \Rightarrow Q)$ and P *implies* Q
- [4] Of course it is possible to represent features with more than two values by Wason cards but the experiment loses its attraction thereby as it is not easy to preserve the overview on the possibilities.
- [5] We are presently testing to guide children to think in terms of binary feature vectors using only two colours, say green and red, and making use of the order of each component of the vector to encode a particular feature. This corresponds exactly to binary encoding of information theory, using 0's and 1's.
- [6] The correct answer is "bottom left" of Figure 12; 83% of the Mundurucu subjects gave this answer, thus exhibiting a high sense for segment proportions.
- [7] Photos 17a and b by courtesy of Elke Kurz-Milcke.