DIETGER HEITELE

AN EPISTEMOLOGICAL VIEW
ON FUNDAMENTAL STOCHASTIC IDEAS

The Woods Hole conference of September 1959 was outstanding of its kind as a meeting of about 35 people interested in education – educationists, psychologists, medical men, and mathematicians such as E. Begle and P. C. Rosenbloom, who discussed improving science education. The results of the meeting were summarised by J. S. Bruner in a chairman’s Report, which, impregnated by his own ideas, evolved into his booklet The Process of Education. In the last decade this work has strongly influenced curriculum development, in particular, in mathematics. Bruner advanced the following theses:

1. The decisive principle of instruction in a subject matter is the transmission of fundamental ideas.

2. "The hypothesis that any subject can be taught effectively in some intellectually honest form to any child at any stage of development" implies that the fundamental ideas are necessary as a guide from kindergarten to university in order to guarantee a certain continuity.

3. Fundamental ideas and concepts will be dealt with on different cognitive and linguistic levels along a curriculum spiral.

4. The transition to a higher cognitive level is facilitated if the underlying subject matter has been prepared in a suitable representation in earlier cognitive stages. In particular the intuitive comprehension of concrete relationships shall be cultivated in the elementary school as long as the child cannot grasp them in a more elaborated analytic form – the prefiguration principle.

Bruner’s contribution seems to me of a particular interest for the instruction in stochastics, which is now entering our schools.

On the one hand advocates of this subject matter are advancing its fundamental (or central) ideas, on the other hand stress is laid on the importance to tie instruction in stochastics to intuitive experiences. Both points, however, are rarely elaborated or made more concrete. In particular the following questions deserve attention:

(a) What would a list of fundamental ideas of stochastical concepts look like?

(b) Why should intuition mean so much for stochastics?

(c) What does "(stochastic) intuition" mean?

(d) How does it develop, and how can it be improved?
In the following we will advance some ideas on (a) and will also touch on the other points.

Proposing a list of fundamental stochastic ideas is a venture since such lists heavily depend on viewpoints, which may widely differ. Arguing from the subject matter one would get another list than if the aim is developmental psychology of random and probabilistic concepts or if instruction in stochastics is seen by an educationist as a means to prepare a "continuous transition from magic and mythical ways of thought to rational-causal explanations". Moreover it may matter whether a subjectivist or an objectivist point of view is chosen, or whether probability is to be considered as a function on sets or as a means to make statements on propositions and their mutual relations.

The present author finds a first relational frame for his views in Bruner's exposition, which considers ideas as fundamental if they are of a great explaining value. Moreover, in Bruner's famous hypothesis the concept 'subject' does not mean an arbitrary subject matter but something that distinguishes educated people, which interpretation restricts the admissible fundamental ideas even more.

In the following I will choose an epistemological pragmatic point of view in Bruner's sense; by fundamental ideas I will understand those ideas which provide the individual on each level of his development with explanatory models which are as efficient as possible and which differ on the various cognitive levels, not in a structural way, but only by their linguistic form and their levels of elaboration.

For instance, the child is handling a crude but adequate explanatory model if in playing with two dice it assigns a better chance to seven spots than to two spots because of the larger number of sums in which seven occurs. This explanatory model can be acquired in playful activities such as proposed by Varga, Engel, and Winter, where no formal, analytical instruction is involved.

By continuity this comparative explanatory model can be developed in order to arrive at a more quantitative one, where the number of favorable possibilities plays a part.

A still more elaborated explanatory model for the same matter would consist in interpreting the sum of spots as a stochastic variable and in stating that the maximum of the probability distribution of this stochastic variable is attained at seven.

What matters here is the constancy of the structure of the explanatory model. The more intuitive model is a coarser — and thus refinable — version of the more elaborate one. If, however, children explain the larger frequency of 'mixed' in two throws with a fair coin by arguing that after 'head' there
comes more often ‘tail’, and conversely, they use a model of explanation which may satisfy them, but does not allow continuation to a more elaborate stage. The model is supported by intuition though it is an “intuition contra­riée”\(^6\), with no continuity of extension.

Intuitive explanatory models are quite often used in teaching, for instance, if fractions are visualised by circular sectors or combinatoric operations by tree diagrams. But they are just visualisations, and a refuge for teachers who feel they have overestimated the abstractive abilities of their pupils. In Bruner’s conception, however, intuitive explanatory models have two functions:

(a) As crude models in an early stage they have an autonomous explanatory value and are helping the child to understand its environment by its own means, long before it can understand the linguistic complexity and sophistication of the underlying mathematical models in their analytic form.

(b) They are pre-establishing the later analytic knowledge in a way that the teacher on a higher grade level can presuppose a favourable intuitive domain when dealing with combinatoric operations.

The large number of paradoxes in stochastics which can be confusing even for experts\(^7\), show that intuitive pre-establishment is more urgent in stochastics than anywhere else. As Feller has stated\(^8\) adults are still able to train their stochastic intuitions! But on the other hand early acquired inadequate explanatory models can apparently develop into firmly rooted intuitions – “intuitions contrariées”\(^9\) – which are difficult to get rid of and can impede the acquisition of analytic knowledge. A typical phenomenon is the concept of independence. If I may generalise my personal experience I dare say that betting at the roulette table on ‘red’ after a long series of ‘red’ gives an uneasy feeling even to mathematicians who know very well the mathematical concept of independence. This is the more strange since 4–5-year-olds do not bother about it though they may find a long series of ‘red’ funny\(^10\). Some time in their development they will acquire some explanatory model regarding this phenomenon, which later on will condense into such inadequate intuitions as we observe with many gamblers, and which are not very likely to be cleared away.

For this reason it looks tempting to offer children stochastical activities as early as in the pre-operational and concrete-operational stage, while trusting that they will develop into intuitions auxiliaires\(^11\), upon which the more analytical instruction in stochastics in higher grade levels can be built. For a few years Varga\(^12\), Winter\(^13\), and Engel\(^14\) have developed many playful activities which put into practice that background conception. However valuable these contributions may be for the progress of mathematical didactics, it is difficult to restrain the heretical idea that the choice of those activi-
ties in stochastics has been left to chance. To avoid the danger that such activities are offered with no aim or purpose, organizing principles are needed, for instance as fundamental ideas, which pierce through the curriculum spiral like guide ropes, as illustrated by the figure.¹⁵ This figure does not mean that on a higher plane of verbalisation all happens in a formalised way. With Fischbein I see enactive-iconic activities dominate only on the intuitive plane, and symbolic representations on the formal plane¹⁶. This is formulated in another way by the statistician D. W. Müller¹⁷: “Mathematics is an extension of intuition with new tools.”

The idea of the curriculum spiral is often used as an argument at the primary school to introduce playful activities that are said to reflect important topological and group theory principles; in recent times the didactical literature has been flooded by such proposals. This makes it even more urgent to reflect upon what is really ‘fundamental’ as long as probability is not yet firmly established at school.

After these general remarks I will try to focus on stochastics in order to find out its fundamental ideas. I have arrived at my list from four angles:

(1) in the frame of Bruner’s conception,
(2) by studying the results of developmental psychology with respect to stochastic ideas,
(3) by studying the multifarious failures of adults in stochastic situations,
(4) by studying the history of probability.

Ad (1): Most of what is relevant, has already been said. The one thing I like to stress is that it is not my principal aim or problem to structure probabilistic instruction in a good deductive order. In stochastics the difficulties do not arise in mathematics but in the applications¹⁸. With Dinges¹⁰
I consider it "as the most important objective of instruction in stochastics that the pupil can deal more safely with scientific pretensions of statistical statements in everyday life." This, indeed, is part of what Bruner calls culture.

Besides this, my first vision angle includes the subject matter. It should be avoided that didacticians try to elementarise contents which are judged relatively unimportant by leading stochasticians. Ad (2): Certainly modern mathematical didactics does not cultivate any more a rigid Piagetian view on a cognitive development that is running off as regularly and automatically as a clock, and certainly it looks for possibilities of acceleration; nevertheless the value of research in developmental psychology has still not diminished. Though it is not a source from which can be deduced which activities are optimal for the learning process (see Aebli) it is a platform from which one can venture into instruction.

Ad (3): While (2) stresses the start, (3) is rather related to the end of a long development. Piaget's thesis that in a strict sense the intellectual development cannot go astray is not possibly true in the domain of stochastics. This is proven by many examples, where adults, even with a college education, behave not much wiser than children. There are fundamental ideas, as there are fundamental errors, and both are counterparts of each other. Such errors bridge the centuries, the ages and the cultural layers, and may be criteria of what is really 'fundamental'.

Ad (4): One may doubt the truth and the applicability of Haeckel's thesis that the ontogenesis is a replica of the phylogensis and still recognise far-reaching analogies between both courses. This is an idea that runs right through Piaget's work but also mathematicians like Thom use it in their argumentation, though for Thom it is a model idea that fits one domain better than others. In the same sense I will use the working hypothesis that the ideas underlying the historical progress of stochastics might be relevant for didactic approaches too.

In my view the relation of model and reality is one of the basic ideas for mathematics in general, and stochastics in particular. It is at present generally agreed upon that a mathematical theory cannot represent the reality completely. "Theories can represent reality only locally in the way a circle and its tangent coincide locally". Laws of nature can be interpreted as hypotheses in the frame of models, where the model is something that replaces reality as operational substrate. This means that models, rather than being abstracted from reality, are imposed upon reality. For this reason a statement within a model is not a statement on reality but can become so by interpretation.

This can fairly well be illustrated by the example of geometry as shown in Scheme A.
The feature that geometry belongs as well to the realm of thought as to the exterior reality, is important in itself but it does not help the pupil to apply his geometry in a more efficient way. Interpreting a mathematical point as a physical point is too natural to be doubted. But transferred to stochastics the scheme looks a bit different, as Schema B.

This scheme may be illustrated by the following example:30

The following game, sometimes called “chuck a luck”, is often played
at small carnivals. The player pays a nickel to play. Three dice are rolled. If any 6's appear, the player gets back his nickel, plus one nickel for each 6 that appears. Players usually believe the bet is favourable and they argue with the following explanatory model: they are right to reason that \[ P(\text{first}=6) = P(\text{second}=6) = P(\text{third}=6) = \frac{1}{6}, \] because of the symmetry of the dice, but they conclude that the probability to get at least one six is \[ \frac{1}{6} + \frac{1}{6} + \frac{1}{6}, \] and on this basis they predict relative frequencies and gains: "in about 50\% of the games I get back the stake but beyond this I have the chance to throw two or more sixes in a single game." If on the basis of this prediction they accept the wager, they are testing their model, and they will experience that in the long run they will lose. This may lead to modifying the model and reach the more adequate conclusion that the gain expectancy of the return is negative.

What seems problematic in the scheme is the occurrence of the concepts of relative frequency and symmetry in the lefthand column. Sometimes it is objected that the concept of symmetry is already a model and registering relative frequencies means quantifying reality, which likewise presupposes model concepts. This seeming inconsistency dissolves as soon as the relations between reality and mathematics are seen in a layer structure. No doubt the numerical data of descriptive statistics can be considered as data of a model. From a higher level one can see them as part of reality, which means passing from the palpable and visible reality to a "new reality" where the relative frequencies emerge as real facts. Symmetry is a similar case.

This is the meaning of the scheme, according to which relative frequencies and subjective degrees of belief are mapped on probabilities, conceived by human minds within a mathematical model. Not until recent times has the fact that probabilities are model data rather than "beings in itself", satisfactorily been understood. Many absurdities have been written by people who intermixed rather than clearly separated the two domains "reality" and "mathematical model".

In early instruction in stochastics this separation can be meaningful, for instance in the standard example of the two dice. It cannot be proved by mathematics that the sample space of this chance experiment should be the set of ordered pairs from \( \{1, \ldots, 6\} \), and that the 36 elementary events should be given a probability of \( \frac{1}{36} \). Up to consistency conditions mathematics does not care which models are chosen. The above standard choice can be heuristically motivated, but the real foundation is the correlate of conclusions from the model in reality (the relative frequencies).

D. W. Müller has called this view rationalisation in aloofness (distanzierte Rationalität): "... an attitude towards factualities... that is not determined by the belief in intrinsic laws but is characterised by an approach with
rational designs such as models, hypotheses, working hypotheses, definitions, conclusions, alternatives, analogies, that is, with the reservation of partial, provisional, approximate knowledge."

If the postulate of rationalisation in aloofness is accepted in teaching stochastics, its consequences should also be accepted. This excludes defining, as is done in some textbooks, concepts as chance experiment mathematically, as well as defining probability in the Laplace or in the von Mises way. Such mock definitions do not contribute to the pupil's apprehension. They rather mean obliterating the frontiers between reality and mathematics. The pupil gets a wrong idea of how mathematics is applied.

This supreme leading view of the rationalisation in aloofness involves that in the lesson chance experiments should not elicit the question "what will happen?" but "what may be expected?" or rather "what do you expect?"

I. NORMING THE EXPRESSIONS OF OUR BELIEF

In my view the first fundamental idea of stochastics is that our intuitive beliefs as expressed in everyday language by formulations like "I believe so", "rather certain", and so on, is normed in a way that impossible events get the probability 0 and certain events the probability 1, and that the relation of "more probable than" is translated into the "\( \geq \)"-relation between real numbers. This means coarsening the complex world around, mapping its multidimensionality upon the one-dimensional real unit interval. This, indeed, is a fundamental idea. Only by suitable coarsening the world becomes accessible to mathematical devices. In this first idea the leading view becomes visible in its aspect that man imposes mathematical entities—numbers—upon reality. Varga has transformed this idea in the didactically suitable tool of the probability scale.

To be sure, if we include all processes which can lead to such numbers (a priori probabilities according to Laplace, a posteriori probabilities) the whole theory of probability is contained in this first idea. But this is not meant here. We only mean scaling the degrees of individual belief in a way as we do with such not quantified notions as temperature.

II. THE PROBABILITY FIELD

Not less fundamental in my view is the idea, first carried out by Kolmogorov, to assign a sample space of observable outcomes to chance experiments, and a \( \sigma \)-field of sets to the field of observable events. This is fundamental not because this, together with accepting an additive normed measure function, led to axiomatising probability in a satisfactory way and thus solving Hilbert's
sixth problem\textsuperscript{34}, but rather because imposing such a model makes the chance experiment surveyable. Winter\textsuperscript{35} has shown how from the elementary level onwards, without any quantitative probability concept, by means of a mere comparative probability fields problems on betting can be dealt with. Tree diagrams, which are undoubtedly an important tool, can be apprehended as structured sample spaces in an iconic representation. I will come back to this point in combinatories.

This second idea seems to me fundamental in still another respect. In research on the development of the chance concept in developmental psychology\textsuperscript{36} it has often been stated that children, in the same way as savages and superstitious civilised people, are confined to a narrowly understood determinism: they believe in coercion by hidden causes or parameters or by a "deus ex machina". Fischbein\textsuperscript{37} and Cohen\textsuperscript{38} explain this phenomenon of concentrating on the single event rather than on the totality, by the fact that our educational system "systematically favours one of the two aspects necessity and chance."\textsuperscript{39} This is reflected by the attitude of quite a few teachers who teach in a way as though every problem, every question has a well-defined unique answer. Even epistemology knows this attitude -- see for instance, Hartwig's\textsuperscript{40} aetiological principle "equal general causes -- equal sets of possible effects with belonging probabilities", which in his view replaces the classical causality "equal causes -- equal effects".

III. COMBINING PROBABILITIES -- THE ADDITION RULE

It is a general feature in mathematics to build more complex models from simple ones, or to reduce complex models to simple ones, and these procedures are used in stochastics too. Though probabilities can be stated by subjective assessments according to the first two ideas, this procedure fails for more complex sample spaces, such as the triple of dice. The operative idea is here the addition rule, which allows new probabilities to be derived from initial ones within the mathematical model.

Though it is not as important in everyday life as the other ideas, its conclusive force in stochastic situations cannot be neglected. The earlier mentioned 'Chuck-a-luck' is a particularly striking example.

IV. COMBINING PROBABILITIES -- INDEPENDENCE

The feature of mathematics in general, and stochastics in particular, of model composing and decomposing, which was already visible in the third idea, is more clearly distinguished if chance experiments themselves are combined and probabilities are assigned to these multi-level chance experiments. What
is fundamental here, is, first of all, the concept of conditional probability \( p(A|H) \), which in applications is interpreted as the probability for \( A \) after \( H \) has happened, in other words, as a measure of how the degree of our belief is changed by new information. Yet starting from everyday experience the idea of independence is even more fundamental; though mathematically it can be reduced to conditional probability, it certainly deserves an independent analysis. It is a fundamental idea to consider chance experiments with no physical causal bond, as stochastically independent. In fact the idea of independent repeatibility of chance experiments shows better than anything else the discrepancy between the mathematical dealing with a theoretical model and its application to reality. On the one hand the mathematical model of independence, as expressed by the product rule, is easy to be mastered; on the other hand it appears in everyday life that many people, even those with a scientific background who are well acquainted with even more complex mathematical models, are not able to apply the idea of independence in a consequential manner in practical situations. A prominent historical example in this respect is d'Alembert\(^{41}\).

From the soldier in battle who always hides in the most recent shell-hole because a burst at the same spot would be improbable, it is not a long way to people who believe in absolutely safe gambling systems, to gamblers in casinos keeping roulette records, and to State Lottery directions who on the back of the lotto tickets publish statistics of under-represented numbers.

**V. EQUIDISTRIBUTION AND SYMMETRY**

The first four ideas do not lead to indications how to compute probabilities. It is a heuristic idea to discover and use symmetries in the problem situation. For instance in the tossing experiment symmetry means that no face of the die is distinguished above the others. This is taken as an argument to accept equidistribution, and the Laplace rule, with all its well-known consequences. If equidistribution is not evident from the start, it can possibly be constructed by refining the sample space and uncovering its symmetries. It should be stressed that equidistribution, which cannot be separated from statistical symmetry, belongs to the mathematical theory only with respect to its formal appearance; as to its content its place is on the heuristic glacis of probabilistic problem situations. For this reason it is impossible to define what statistical symmetry means. In fact, some principles have been stated, such as that physical symmetry implies statistical symmetry. The insufficiency of this principle is illustrated by the figure of a chance experience with balls, falling along a pipe system\(^{42}\) (see the figure on the next page). Equidistribution is here a reasonable assumption though there is no physical symmetry.
The principle of insufficient reason would cover all cases but it is too vaguely formulated and may lead to paradoxes and inconsistencies, in particular if infinite sample spaces are concerned. It seems that the idea of symmetry and equidistribution cannot be grasped intentionally but only in the long run by working with a multifarious material of examples. In many situations, for instance, if a physician takes a blood sample to be analysed, in the classical problem of two people meeting each other \(^{43}\), the equidistribution is the most natural working hypothesis, the adequacy of which is only \textit{a posteriori} tested. At the start of analytical statistics, in Bayes' theory, it played a paramount part as the \textit{a priori} distribution which is assumed as an equidistribution as long as there is no argument against it.

One of the nicest applications of equidistribution is the story about the student \(^{44}\) who every morning out of two streetcars - two lines with the same interval, one leading to the university, and the other to his girl friend - chose the first; he considered it as his destiny that in nine out of ten cases he arrived at his girl friend's. If this story is true, Laplace's rule could have contributed to rationalising a magic world picture.

\textbf{VI. COMBINATORICS}

It is too simple a policy to consider combinatorics as ancillary to probability as it might seem natural from the viewpoint of mathematical structure. Drawing, for instance, from an urn with four distinguishable objects three objects, is a three steps chance experiment, which is meaningfully interpreted in the sample space of \((4, 3)\)-permutations. This connexion between multi-step chance experiments and permutations becomes clear in a tree diagram.
The tree diagram as an iconic representation is of fundamental importance because it visualises the multi-step structure of the experiment as well as all possible results, and so it works with other combinatorial operations. For this reason the basic combinatoric operations rather than being simply standard algorithms for computing probability fields of complex chance experiments, provide for straightforward insight, in particular in their iconic and enactive form, into the interior structure of chance experiments and the enchainment of successive experiments within a larger complex.

This view is supported by results of developmental psychology and J. Piaget's theories on the development of the chance concept which in their totality, though dating back about twenty years, have not up to now, as far as I know, critically been checked, either by empirical means or in a theoretical approach. One of Piaget's main theses in this work says that the path to comprehension of both the chance and the probability concepts is leading along the basic combinatoric operations.

Beside this it is the power of combinatorics to classify complex chance experiments according to their composition from simpler ones in certain standard types and by this way to provide the active mathematician with standard models for numerical processing. The almost self-evident original subideas are leading to these standard models of permutations, variations, and combinations:

The product rule: if $n$ paths lead from $A$ to $B$ and $m$ paths from $B$ to $C$, then there are $nm$ total paths leading from $A$ to $C$.

The quotient rule: If a set of $n$ elements is partitioned into subsets of $m$ elements, then the number of partition classes is $n/m$.

There are good reasons to believe that the first subidea of "combinatoric counting" is badly needed for the acquisition of a multifariously applicable number concept, even though it does not fit very well into the picture of an
elementary school mathematics where the cardinal aspect is wrongly stressed. Even very young children, 5-6-year-olds, can grasp this basic counting principle in concrete examples and transfer it successfully.

VII. URN MODEL AND SIMULATION

In principle it is possible to assign urn models to the greater part of chance experiments, at least to those with a countable sample space. The urn idea looks fundamental for various reasons.

First of all stochastics knows quite a few concepts that are firmly rooted but defy any rigorous definition. A typical example is "random choice". It can be tested a posteriori whether a sample may be considered as "random" but the process of getting random samples is by principle inaccessible to mathematical definitions. The only way to describe it – and a highly efficient one – is concretising it by the urn model. Secondly, urn and chance experiments can be compounded into new ones, so-called hyper-urns corresponding to the compounded experiment. This allows one, as Polya has shown, to simulate so complex chance processes as the course of the weather in a most striking way by a sequence of urns. The cue word "simulation", which in the statistical jargon means something like isomorphism, indicates a third reason for the central importance of the urn model.

In certain contexts it seems reasonable to consider two urns as isomorphic if corresponding ratios are equal. By the relation between urns and chance experiments this concept of isomorphism can be transferred to chance experiments: chance experiments are isomorphic if the corresponding urns are so. This implies the possibility to realise chance experiments by means of urns, tops, random numbers, and so on, that is, by other chance instruments – this is the meaning of the Monte Carlo method. To be sure, if Sobol's definition is accepted "the Monte Carlo method is a numerical method of solving mathematical problems by means of modelling with stochastic magnitudes", another aspect is meant, the idea of stochastic variable.

VIII. THE IDEA OF STOCHASTIC VARIABLE

If there has been one conception that lifted probability above the level of mathematisation of gambling experiences, then it might safely be said that the conception of stochastic variables and the conceptual inventory related to it have been responsible for the multifarious applications of applied statistics. There are compelling reasons why mathematicians of past centuries who did not know the concept of stochastic variables, got into serious trouble with such problems as the Petersburg paradox, and why James Bernoulli
needed twenty years to discover and to prove his weak law of large numbers\textsuperscript{53} – problems that are dealt with in the present language of stochastic variables in a few lines. Not only for theoretical probability was the idea of stochastic variable fundamental, in everyday life, too, stochastic variables are important, in hazard games, in queueing problems, and in many physical processes. On the strength of multifarious – albeit unconscious – experiences with stochastic variables in everyday life, one can agree with the view that the intuition of magnitudes that are delivered with participation of chance, arises even earlier than that of chance experiment\textsuperscript{53}. As explanatory model the concept of stochastic variable plays an important part in three respects: the distribution of a stochastic variable, its expectancy, and compounding stochastic variables to get new ones.

For little children all stochastic variables are in principle equidistributed; it looks like a moral principle: everybody should get his part\textsuperscript{54} Even older children are inclined to assume the equidistribution in such simple experiments as two coins: head-head, tail-tail, mixed. This suggests the question whether the stochastic variables should be approached via the special case of equidistribution, or rather, according to Dienes’ deep-end-principle via general distributions. Whatever the answer may be, it cannot be denied that beside the equidistribution the normal distribution plays a fundamental part in the explanation of the world around us. A convincing explanatory model for this fact is delivered by the central limit theorem, which, however, is not accessible to a deductive approach on any level below the university. This does not exclude considering whether some intuitive grip on, and experimental-inductive approach to this theorem should not be counted to the stock of knowledge of every educated man, and therefore be considered as a “subject” in Bruner’s sense.

The most important features of a distribution function are its expectancy and its standard deviation. Both of them are fundamental for man in our society. They should enable him to face statistical data safely and critically. In the long run we cannot be content with feeding people with averages, without ever mentioning measures of dispersion as is the habit of our media.

A special chapter is the expectancy of a stochastic variable. It plays a purely biological part in the probabilistic approach to theories of learning where it appears that even not consciously thinking animals in Skinner boxes can learn expectancies\textsuperscript{55}. On a higher level, for consciously thinking humans, it possesses a great explanatory value; intuitively it is most often interpreted as the arithmetical mean of values of a stochastic variable, obtained if the basic chance experiment is often enough repeated under “identical conditions”\textsuperscript{5}. This intuitive comprehension is made more precise by the laws of large numbers.
IX. THE LAWS OF LARGE NUMBERS

One should distinguish between an empirical law of large numbers, or "principle of large numbers", and purely internally mathematical laws of large numbers. The first is straightforwardly observable in the reality, for instance, in the well-known examples of raindrops on paving tiles. Rain is a typical random mass phenomenon where individual events are unpredictable in principle. It is philosophically interesting that globally a regularity arises which seems inherent to the course of nature - Wagemann\textsuperscript{56} called it in a pregnant way "individual liberty under collective constraint" - which cannot possibly be explained mathematically, but what really matters is that this principle of large numbers possesses an internally mathematical correlative in the laws of large numbers which can be derived from the model of probability field so as to justify it by this consequence as a good model. The sometimes expressed opinion that the laws of large numbers should be measuring prescriptions for unknown probabilities or that the parameter $p$ in Bernoulli's law

$$\lim P \{|1/n(X_1 + \ldots + X_n) - p| > \varepsilon\} = 0$$

should be the "genuine" probability, seems to me inadequate, because of its intermixing of "factual and design plane".

I do not dare answer the question whether the postulate of separation of reality and model\textsuperscript{57}, or even of consciousness about it - the rationalisation\textsuperscript{58} in aloofness can be possible and effective on every cognitive developmental level, but I think it would be worthwhile leading the individual to very early empirical experiences of this phenomenon of "individual liberty under collective constraint". In particular, the partial aspect of individual liberty gives troubles not only to children but also to many adults, as appears not only from Piaget's and Cohen's\textsuperscript{59} results, but also from experiences in casinos and bizarre stories in scientific treatises in the past\textsuperscript{60}. In the main two inadequate models are again and again used by adults too:

The alternation model: "after 'head', 'tail' is more likely than 'head'", or, "after a strain of bad luck the tide will turn", the seriation model: "if it is a strain of luck, I do not get out".

In uncritical papers in the media\textsuperscript{61} the second model occurs as a mysterious "law of the series"; it seems to be basic to many kinds of superstition. In both models the belief in the existence of winning strategies and the lack of understanding independence are involved.

The didactic opportunities of empirical experiences in this ninth idea are more restricted than school texts would make us believe. Random sequences in the class room converge slowly\textsuperscript{62}, and possibly not at all when they
are needed in demonstrations. It seems a better opportunity to build such series into games. In fact "people do not like to learn, but they do like to gamble".63.

X. THE IDEA OF SAMPLE

The example of the raindrops on the pavement show a tenth idea, that of sampling. It is fundamental not only in the show examples - examples of psychological investigations, medical diagnostics, industrial quality control - but all of our knowledge and judgment is based on samples. Prejudgment is nothing else but judgment on the strength of non-representative samples.

Since thinking, judging, inferring is only possible on the basis of samples, it matters to have people arguing cautiously and critically. Like the professional statistician everybody should consider sampling and its consequences as crude models to explain the reality, and, clearly understand in every particular case that his conclusions are statistical, and moreover, which would be their consequences, and what harm a wrong decision could do. Children at the elementary school level can grasp that decisions may depend on the accident of choice, and that a good choice is an ideal case, if it is suggested by appropriate activities like the following:64

The teacher shows a bagful of rice grains. How many are there? Can we tell it without counting? The grains are scattered over a square lattice and shown by overhead projector. This impulse is strong enough to get children starting. Some squares seem more appropriate than others - it is obvious how this can be continued. The reason why this special activity works very well is the simultaneous visual exposition of both the totality and of many samples. In general, activities seem to promise more success if they are embedded into geometrical situations.

On higher grade levels examples with a less direct connexion to geometry can be chosen65: The pupils are given the problem of how many children does the average family have. The pupils are inclined to choose their own class as a sample and evaluate these data. In the course of the lesson it occurs to them that this sample cannot be representative since families with no child are lacking in the sample (and with a larger number of children are over-represented). This is a fundamental insight.

FINAL REMARKS

My list of fundamental ideas is a sample of stochastic models, prescriptions, and factualities. Is it representative?

As announced in the beginning, this depends on the viewpoint. My starting point was the question: Which stochastic situations will the individual pre-
ponderantly meet, and by which explanatory models can they be covered? This is a frame, and within the frame structurations were tried, concentrations in the form of general ideas were pursued. The resulting list is something like a model, though not for solving stochastic problems but for building connected curricula of stochastic. The usefulness of such a model can only be shown by using it in teaching at all levels. The major task of didactics of probability in the next few years is not writing deductively well structured chapters on stochastic for grade level $n$ – probability as an abstract mathematical domain --, or to produce still more play activities in no systematic setting. Rather than this it is necessary to integrate stochastic activities as early as possible with activities in arithmetics and, even more, in geometry, and in all cases to respect and to develop meaningful connexions with reality, with the world of the pupil. For this purpose we need teachers who know what is really fundamental in stochastic. For instance, it would be a bad thing if our elementary teachers would teach elementary stochastic as a field of 'applications' of the so-called set theory. Exactly this might happen as our elementary teachers in Germany are taught much 'set theory', but little, if any, stochastic.

NOTES

4 Cp. first paragraph (2).
9 Cp. Note 6, p. 265.
10 Piaget, J. and Inhelder, B., *La genèse de l'idée du hasard chez l'enfant*, PUF Paris 1951, p. 110–111. (Piaget's experiment is not identical, but isomorphic with the present one.)
11 Cp. Note 6, p. 265.
204 DIETGER HEITEL


For instance, people active in instruction, seem to believe that Kolmogorov’s axiomatic system has contributed in a decisive way to the progress of probability and statistics, apparently because they endeavour to see mathematics from the viewpoint of mathematical structures. Statisticians, however, and even probabilists view Kolmogorov’s system rather as a finishing touch than as progress in essentials.

21 Cp. Note 16, p. 293.


25 It would certainly be impossible and ineffective to imitate all errors and deadlocks.

26 Cp. Note 16, p. 293.


28 According to Hinderer, K., Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin 1972, p. 1 sq. (Hinderer does not give an iconic scheme).


30 In the classroom this example causes trouble, for instance if the sum of points is considered as a stochastic variable: The pupils argue on the strength of commutativity of addition that (2,5) and (5,2) need not be distinguished.

31 Cp. note 17, p. 167.

32 Cp. note 12, p. 352 sq.

33 Cp. note 16, p. 293.


35 It would certainly be impossible and ineffective to imitate all errors and deadlocks.

36 Cp. Note 10. See also Cohen, J. and Hansel, M., Glück und Risiko, Frankfurt 1961 (1955()).


38 According to Hinderer, K., Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin 1972, p. 1 sq. (Hinderer does not give an iconic scheme).


40 In the classroom this example causes trouble, for instance if the sum of points is considered as a stochastic variable: The pupils argue on the strength of commutativity of addition that (2,5) and (5,2) need not be distinguished.

41 Cp. note 17, p. 167.

42 Cp. note 12, p. 352 sq.


44 Winter, H. and Ziegler, T., Neue Mathematik Bd. 2, Dortmund 1971, p. 112.

45 Cp. Note 10. See also Cohen, J. and Hansel, M., Glück und Risiko, Frankfurt 1961 (1955()).


48 Cp. note 17, p. 167.

49 Cp. note 12, p. 352 sq.


51 Winter, H. and Ziegler, T., Neue Mathematik Bd. 2, Dortmund 1971, p. 112.

52 Cp. Note 10. See also Cohen, J. and Hansel, M., Glück und Risiko, Frankfurt 1961 (1955()).


55 Cp. note 17, p. 167.

56 Cp. note 12, p. 352 sq.

57 Cp. note 16, p. 293.


62 See the example in Freudenthal, Cp. Note 18, p. 190.

63 Freudenthal, H., ‘General ideas by comprehension or apprehension’, Lecture at the International Colloquium on theoretical problems of mathematical instruction at the primary level.


52 Cp. Menges, G., Note 43, p. 11.


54 Cp. Note 10.


56 Cp. Wagemann, E., Note 24, p. 22.

57 See Freudenthal, Note 18, p. 583.


59 Cp. Note 36.


61 Cp. _Quick of 7 March 1974: “Ein Unglück kommt selten allein. (disasters agglomerate – a German proverb, which actually means ‘it never rains but it pours’). Millions say it without knowing how true it is. Often catastrophes, accidents, diseases pile up in a way which cannot be explained by mere chance. Scientists call it the magical law of the series.”_


64 Personal communication by G. Müller, Dortmund.

65 From H. Freudenthal (from a lecture).