EVOLUTION OF FORMS OF REPRESENTATION IN A MODELLING ACTIVITY: A CASE STUDY

Rossella Garuti Carlo Dapueto, Paolo Boero

I.R.R.E. Emilia-Romagna Dipartimento di Matematica, Università di Genova

The report describes a mathematical modelling activity of a natural phenomenon (transmission of hereditary characters in a codominance case) using the concept of model (as represented by the diagram in Fig. 1) as a theoretical instrument. The chosen tool enables us to show how the construction of a link between reality and a model is related to the evolution of the graphical representations adopted by the students. The lack of such evolution may be either an obstacle to the modelling activity or favour an inappropriate adoption of models generally used for other phenomena. Our analysis also compares the resolution processes enacted by students belonging to different educational environments.

THEORETICAL FRAMEWORK

Most research on probabilistic thinking concerns students' evaluation of a stochastic situation (prediction or interpretation of outcomes); few research studies concern probabilistic modelling, i.e. the construction of a model to interpret stochastic real world phenomena (for a general survey, see Borovcinik & Peard, 1996). As concerns semiotic aspects, Pesci (1994) and Dupuis & Rosset Bert (1996; 1997) referred (resp.) to Fischbein (1987) and Duval (1995) for general theoretical frameworks. They pointed out the crucial role of semiotic tools in probabilistic thinking, but did not deal with modelling of complex phenomena. The process of progressive schematisation of a phenomenon in order to get an interpretation in terms of a probabilistic model still needs to be carefully investigated. In particular it would be necessary to consider the nature of some specific semiotic tools (tree graphs, double entry tables, etc.) and the relationships that each of them allows to establish between the phenomenon on one side, and probabilistic thinking on the other.

In Dapueto & Parenti, 1999, the epistemological concept of model, represented in the diagram in Fig. 1, is discussed (in comparison with preceding studies: see Blum& Niss, 1991; Norman, 1993) and used in order to tackle some problems of situated teaching-learning. They describe the diagram with an example: *"Let R be a quadrangular field that we wish to evaluate economically an AR the extent of that field. If we note that the field is more or less rectangular, we can consider the lengths of two consecutive sides as the factors ER that determine* *the extent*. […] *Let us associate the lengths ER to two numbers*, a *and* b, *which express the lengths in a particular unity* U;[…] *Let us associate to AR* a*b*; this is AM*. […] *Hence, the model M is the multiplication (or rather the formula* S=a*b, *if* S *expresses the extent in square whose side is* U*)*". Subsequently the authors write *"in mathematical modelling (i.e. when artefacts* [used to build the model] *are mathematical objects) the passage from ER to EM is called mathematization. But also the use of mathematical artefacts in building a physical, biological,…model (for instance a physical law or a bionic model)l is a mathematization"*

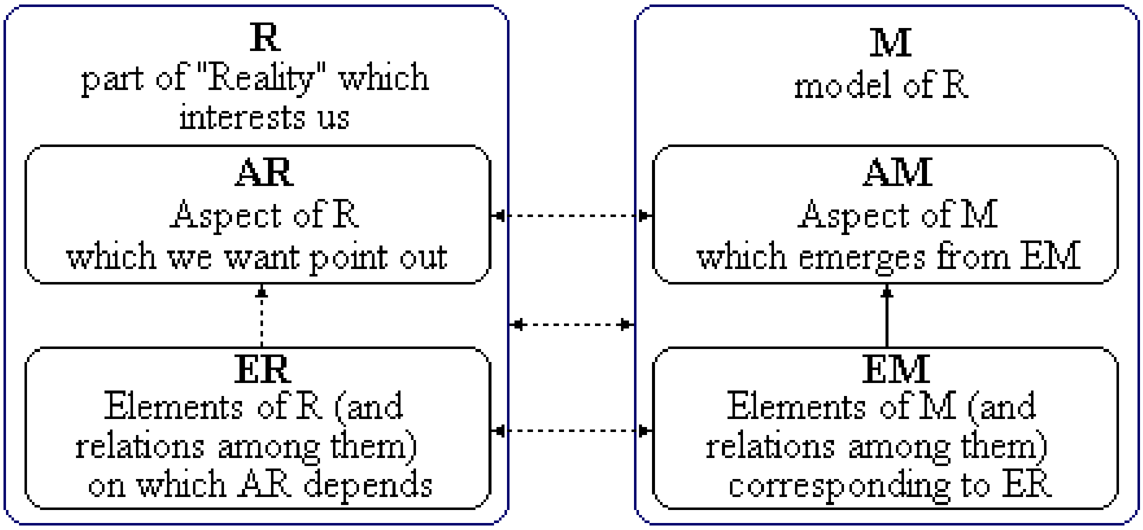

Fig.1

We used this representation of the concept of model in order to carry out a detailed analysis of VIII grade students' behaviour while performing a complex activity of mathematical modelling of a natural phenomenon, i.e. the transmission of hereditary characters in a codominance case. It enabled us to study how the construction of a link between reality and a probabilistic model is influenced by the evolution of the representations adopted by the students, particularly in passing from ER to EM (the representation of the phenomenon through schemata, graphs, symbols, concepts drawing on different branches of learning). The lack of such evolution can either hinder the modelling activity or induce an improper borrowing of representations generally used to analyse other phenomena.

THE TEACHING EXPERIMENT

The teaching experiment involved two VIII grade classes: the first located in the North of Italy with 20 students, the second located in Cataluña (Spain) with 26 students. It is important to remember that, in Italy, in the lower secondary school, Mathematics and Sciences are taught by the same teacher; in this case the teacher has been the same since grade VI. In the Spanish classroom the situation was the same. Both class teachers belong to the Genoa Research Group in Mathematics Education. In both classes the didactical contract implied the argued production of hypotheses and the comparison of such hypotheses through discussion. Particularly the students were familiar with various experiences of mathematical modelling of physical and natural phenomena. (see Boero & Garuti, 1994 and Boero et al.,1995) In the curriculum of the Genoa Group Project the study of Genetics represents the core of Sciences teaching in seventh grade and it is connected to Mathematics through the introduction of the probability model. At the beginning of the activity great attention is given to the students' conceptions on transmission of hereditary characters, and Mendel's Laws are reconstructed through the guided reading of the essay presented by Mendel to the Naturalistic Society in Brno in 1865 (original title: *Versuche über Pflanzen-hybriden*). Teachers greatly emphasise the fact that Mendel (differently from previous scientists who had already recognised phenomena of dominance and segregation) did not limit himself to a qualitative description of the problem but, by means of mathematical concepts, developed a theory that allowed him both to quantitatively describe the problem and to deepen the biological interpretations. In particular the relationship between the analysis of the experimental frequencies and the probabilistic model is carried out, referring both to the data collected by Mendel and to the flipping of two identical coins, as representation and simulation of the phenomenon (the genes become the sides of the coin) and as prototype for favouring the students' conceptualisation. The only difference between the two classes refers to the forms of representation used: the Italian students developed their own form of representation (essentially tree graphs, keeping always a strong link with a more iconic representation of the phenotype) and ended up in the use of pairs of letters to denote the genotype; while the Spanish students learned to use double entry tables, the so called "Punnett square", to represent the possible combinations of genes.

The task: a priori analysis

The Mendel's surprise

The Mirabilis Jalapa is an ornamental plant of which two varieties are known, one producing only white flowers and the other only red flowers. So this plant gave Mendel another excellent chance to study the effects of crossing over an isolated and flashy characteristic such as colour. Mendel prepares two groups of plants: one with red flowers only and the other with white flowers only. Then he crosses the two types of plants, waiting for seeds production. The next season all the seeds are planted and he waits for the plants to grow and for the flowers to bloom. At the blooming Mendel finds out that all the plants are covered with pink flowers. The original characters, white and red colour, had disappeared and seemed to have mixed up like water and wine. Like in the previous crossing *studies Mendel decided to cross the pink plants. The following year some of the new produced plants carry pink flowers, some red and some white flowers.*

1) *When Mendel realises that the second generation produces pink, white and red flowers, he is able to calculate the percentage of each colour. According to your opinion, which are the percentages forecast by Mendel and why?*
2) *Supposing you were Mendel, how would you interpret the observed results of the crossing?*

In the classical case studied by Mendel, the *Pisum sativum* case, one character is dominant on the other and after the crossing of two pure lines, for instance yellow seeds peas (YY) and green seeds peas (GG), in the next generation the so called hybrids (YG) show only yellow seeds plants, and in the subsequent generation the probability to obtain yellow seeds plants is 75% (three possible combination out of four: YY, YG, GY, while the probability to obtain green seeds plants is 25 %, one combination out of four, GG). This makes sense provided that the distribution is uniform, i.e. no gene is 'favoured' with respect to the other (in the case of the coins this corresponds to the hypothesis of their equality).

The case of *Mirabilis jalapa* represents a case which Mendel's first law cannot be applied (Law of dominance: *In a cross of parents that are pure for contrasting traits, only one form of the trait will appear in the next generation. Offspring that are hybrids for a trait will have only the dominant trait in the phenotype*). To correctly understand the results, it is necessary to switch from this model at the phenotype level, i.e. the shown characteristics, to that at the genotype level, i.e. the genes combinations. In this way the dominance can be singled out as the hypothesis subjected to the applicability of the law. An interpretation of the new phenomenon can be built through a new process of modelling: starting from the genotype level; finding an explanation of the fact that the hybrid is not similar to either parents, but shows a new character, the pink colour of the flower; and producing a probabilistic evaluation for the distribution of characters in the next generation.

The use of the previously described diagram (Fig.1) can help to better recognise different steps in the representation and interpretation of the phenomenon. In our example R (the part of reality which interests us) is the transmission of the flower colour, AR (the aspect of R, which we point out) represents the way in which the flower colour depends on the colour of the flowers of the previous generation, and finally ER are the colours of the plants of the different generations. All these elements can be obtained from the verbal description of the problem. At this point the crucial step in the modelling activity consists of the representation of ER through pairs of letters (W and R), which symbolise the genes that control the colour of these flowers. The representation of the situation AM is then realised highlighting the possible coupling either through a tree graph (with the edges labelled by the relative probabilities) or with a double entry table (under the hypothesis that all cells have the same probability). Referring to this representation the model M can be described through a formulation in probabilistic terms (in the second generation a flower has a 50% probability to be pink, 25% to be red and 25% to be white).

To be able to interpret the phenomenon we need, therefore, to go through the building up of a representation of the different elements (and of their respective relationships) which characterise the phenomenon. This representation needs to be abstract enough to facilitate the analysis of the phenomenon, but at the same time, it needs to remain context-related in order to allow a natural managing of the interactions with the phenomenon that are necessary to the elaboration, discussion and realisation of the model and its possible revision. This is a fundamental aspect from a didactical point of view. The problem of the *Mirabilis jalapa*, proposed to the students one year after studying Genetics in VII grade, represents a challenge for them. The results of the crossings seem to be in contrast with Mendel's hypothesis: it is a case of codominance, that the pupils have never met before, in which the presence of two different genes determine an aspect of the character different from that of the parents. Furthermore, the fact that the pink colour can be interpreted as a mixture between white and red can make some students reconsider those pre-mendelian hypotheses strongly present in the class before introducing Genetics.

SOME OUTCOMES

An overall view of students' behaviour

a) In the Italian class 13 out of 20 students produced a correct modelling of the situation and answered the first question (regarding the probability of the different colours). Only 4 students were able to formulate an explicit codominance hypothesis, the others limited themselves to describe the situation, without working out any conclusion.

b) In the Spanish class 8 out of 26 students produced a correct modelling of the situation and answered the first question; in spite of this none of them suggested the idea of codominance.

Italian students' behaviour and evolution of their representations

All the students describe the situation through graphical representations or drawing flowers, or simply writing the names of the colours, reaching a schematic representation of the crossing, always at the phenotype level. They translate into images what is described in words in the text. They still are at ER level and they should pass to EM level. At this point the students' behaviour differentiate. Some students (7 out of 20) stop here (Fig. 2) and try to answer the first question, in some cases applying a probabilistic reasoning. As an example

Giulia writes: “*Mendel will find equal percentages for pink, red and white flowers, because the three colours have the same probability”*. She extends the uniform distribution hypothesis to the colours, while this hypothesis is only valid for the pairs of genes. In other cases the answers are not connected to any probabilistic reasoning, but to personal conceptions. Sara writes: “*I think that Mendel foresees this percentages: 25% pink flowers, and the remaining white and red, because pink is not pure, so it is probable that in the first experiment the pink flowers are less than those of pure colours*”. In this case the student is not guided by any model, but by an idea of purity that prevails. These students are not able to interpret the results of these crossings as Mendel did and only one student from this group shows surprise. Luca: “*If I were Mendel I wouldn’t be surprised of the absence of white and red, but certainly more surprised of the coming out of pink colour. According to Mendel the characters are inherited from father or mother. By the way, the mixture theory came out during discussions in class. Often it happened to hear that in a family the father was dark-haired, the mother blond and the son brown-haired. But this never came out in Mendel. The explanation that I could give to myself if I were Mendel is that since the genes are both present, they appear mixed up*”.

The rest of the students (13 out of 20), after describing the situation exactly in the same way as their classmates, associate a pair of letters, representing the pair of genes, to the flower colour and this allows them to connect themselves to the learnt theory, to find out the possible combinations and to calculate their probability (2/4, 1/4, 1/4 or 50%, 25%, 25%). The transition from an iconic representation (still at ER level) to the one using letters, consistent with the studied theory, (we are now at EM level), allows them to recognise the pink flower as a hybrid and to correctly calculate the percentage. (Fig.3). Andrea writes: “*The percentages are 50% pink, 25% red, 25% white. It’s what Mendel thought because the pink flowers are normal hybrids, even if the colours are mixed up to form the pink colour (RW), the genes are equal to any hybrid. When crossing to different pure lines (WW and RR), I obtain a hybrid (RW). Then, crossing the hybrids, I will get two hybrids, one pure white and one pure red, out of four".*

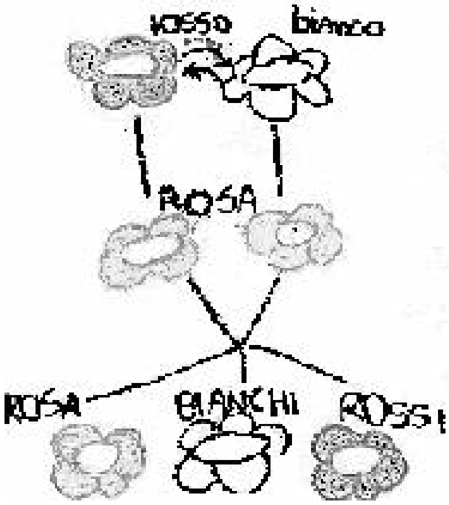

Fig.2

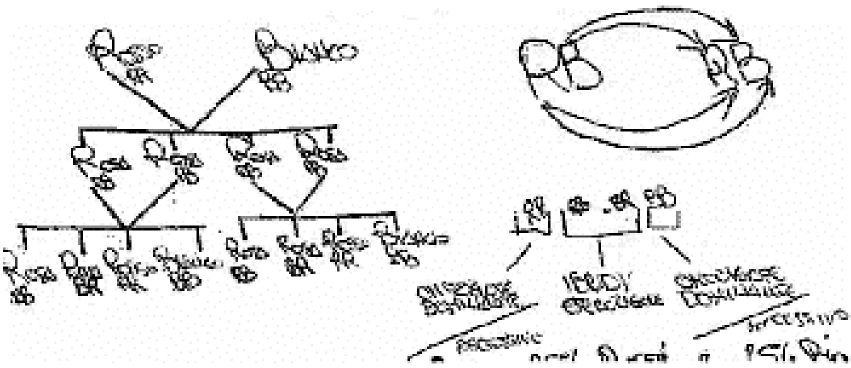

Fig. 3

The second question concerns the interpretation of the results of the crossing of the *Mirabilis jalapa* according to Mendel's theory. In order to be able to do so it is necessary not only to produce a model for the *Mirabilis jalapa* (passing from ER to EM), but also to model the case studied by Mendel (*Pisum sativum*) and compare the two. They are indeed different realities that can be interpreted by the same model, but the different results need to be interpreted (probability distribution 1/4, 2/4, 1/4 vs 1/4, 3/4). In the frame of our theoretical tool we need to pass from AM to AR. Only four students out of the thirteen of this group are able to formulate a codominance hypothesis and, again, the representation helps them. These students produce a graphic representation also for the known case of *Pisum sativum* and compare the two representations, realising in this way that those elements, showing in the case of *Pisum sativum* the dominant character with a 75% frequency, correspond to the 50% of pink flowers added up to the 25% of one of the other two colours in the case of *Mirabilis jalapa*. This observation allows them to make the codominance hypothesis. As an example Laura writes: "*In his previous experiments Mendel obtained 25% of recessive and 75% of dominants, but included in that 75% there was also a 50% of hybrids not visible. In this case the hybrids show not the dominant colour but the pink colour It seems that white and red can show up together, so there is* *neither dominant nor recessive, and this means that you can obtain not two but three types* (phenotypes). *In this way it is easier to understand what is hybrid and what is pure*”. Davide explains: “*The pink colour is like the red with a "mutation". This "mutation" has been created in the past or it is an effect that happens when the dominant characters are close to the recessive ones*”. The other students do not produce any hypothesis consistent with the results of the crossing and none of them produces a representation of Mendel’s experiment to be compared with the case under study.

Comparison with the Spanish students

We recall that Spanish students had learnt to use the double entry table in order to represent the results of the possible crossing. Twelve out of 26 students do not utilise the learnt representation in order to model the situation, but apply the probabilistic model to the three different colours of the flowers, giving the same arguments as the Italians students. Six out of 26 students make the hypothesis of a gene controlling the pink colour, without realising that this hypothesis must be logically refused (it is not possible that the first generation plants, the ones with red and white flowers, carry a gene responsible for the pink colour, since they are pure). They pass from ER to EM since they use letters to represent the pairs of genes, they lean on the learnt representation, the Punnett square, but they stop at this level: they are not able to interpret AM, aspect of model, in relation to the real situation. As an example Josè writes: *Surely the second generation parents will be carriers of the pink gene, and that is the reason why the pink flowers come out. Let’s draw a table* (See Fig.4). *I would say that Mendel foresees that 0% will be red, 0% white, 25% pink and 75% carriers of red and white*”. We can see that what the student writes is contradictory with what is written in the text of the task; he is driven by an initial hypothesis, applies a learnt representation, but he does not control at all the meaning of what his modelling says (he cannot say which colours are the flowers of the remaining 75%). 8 out of 26 students give a correct prediction of the percentages of any type of plant, they use the double entry table fluently and establish causal relationships between pairs of genes (genotype) and aspect of characteristics (phenotype), passing from AR to AM and vice-versa. Nevertheless none of the students introduces an interpretation in terms of codominance.

	V	R
B	BV	BR
R	RV	RR

Fig. 4

CONCLUSION

The epistemological concept of model, as represented in Fig. 1, was used to describe the process of progressive schematisation of a complex situation that needed to re-construct the probabilistic interpretation of the phenomenon under scrutiny. The concept was useful both in order to detect the points of that process, where the contact between reality and schematisation was lost by some students, and in order to understand some differences between two classes, where semiotic tools for schematisation had been introduced and used in different ways. Implications for teaching of modelling in the case of probabilistic models concern the need that students learn: to build and use different kinds of graphical representations (from those very near to the phenomenon, to those more suitable for calculation); and to keep under constant control the relationships between such representations and reality.

REFERENCES

Blum,W. & Niss, M.:1991,'Applied mathematical problem solving, modelling, applications and links to other subjects’, *Educ. Studies in Math*, 22, 37-68.

Boero, P. & Garuti, R.: 1994, 'Mathematical modelling of the elongation of a spring: given a double length spring', *Proc. of PME XVIII*, Lisbon, vol.2, pp.384-391.

Boero, P., Dapueto, C.,Ferrari, P.L., Ferrero, E., Garuti, R., Lemut, E., & Parenti, L.:1995 'Aspects of mathematics-culture relationship in mathematics teaching-learning in compulsory school', *Proc. of PME XIX*, Recife, vol.1, pp.151-166.

Borovcinik, M. & Peard, R.: ‘Probability’, in A. Bishop et al. (Eds.), *International Handbook of Mathematics Education*, pp. 239-287, Kluwer A. P., Dordrecht.

Dapueto C. & Parenti L.:1999, Contributions and Obstacles of Contexts in the Development of Mathematical Knowledge, *Educ. Studies in Math.*, 39, 1-21.

Duval, R.: 1995, *Semiosis et pensée humaine*, Peter Lang, Bern.

Fischbein, E.: 1987, *Intuition in Science and Mathematics*, Reidel, Dordrecht.

Dupuis C. & Rousset-Bert S: 1996, 'Arbres et tableaux de probabilitè: analyse en termes de registres de representation' *Reperes-IREM* 22, 51-72

Dupuis C. & Rousset-Bert S.:1997, 'Tree diagrams in probability: a real register of representation', *Proc.of PME XXI*, Lahti, vol.1 pp.230.

Norman, D. A.:1993, *Things that Make us Smart*, Addison Wesley Publishing Company, Addison.

Pesci, A.: 1994, 'Tree graphs: visual aids in casual compound events', *Proc.of PME XVIII*, Lisbon, vol.4, pp.25-32.