

## A quantitative method to analyse an open-ended questionnaire: A case study about the Boltzmann Factor

ONOFRIO ROSARIO BATTAGLIA<sup>(1)</sup> and BENEDETTO DI PAOLA<sup>(2)</sup>

<sup>(1)</sup> *Dipartimento di Fisica e Chimica, Università degli Studi di Palermo - Palermo, Italy*

<sup>(2)</sup> *Dipartimento di Matematica e Informatica, Università degli Studi di Palermo  
Palermo, Italy*

received 26 October 2015

**Summary.** — This paper describes a quantitative method to analyse an open-ended questionnaire. Student responses to a specially designed written questionnaire are quantitatively analysed by not hierarchical clustering called  $k$ -means method. Through this we can characterise behaviour students with respect their expertise to formulate explanations for phenomena or processes and/or use a given model in the different context. The physics topic is about the Boltzmann Factor, which allows the students to have a unifying view of different phenomena in different contexts.

### 1. – Introduction

Extensive qualitative research involving open-ended questionnaires as well as standardized multiple-choice tests provided instructors tools to probe their students' conceptual knowledge of various fields of physics. Many of such studies examined the consistency of students' answers in a variety of situations, where the underlying physical systems are found similar from an expert point of view, and tried to develop more detailed models of consistency in reasoning strategies of the tested student populations or to subdivide a sample of students into intellectually similar subgroups. Bao and Redish (Bao and Redish, 2006) introduced a framework, model analysis, for exploring the structure of the consistency of the application of student ability by separating a group of students into intellectually similar subgroups. Qualitative and quantitative research methods have been applied in order to establish a quantitative representation framework by analysing students' alternative knowledge and the probabilities for students to use such knowledge in a range of equivalent contexts. By integrating qualitative and quantitative methods, results of qualitative research are used as the basis for the theoretical assumptions to be employed in the data analysis in order to evaluate the potential causal pathways for the inferential analysis as well as the issue of context dependence.

The problem of taking a set of data and separating it into subgroups where the members of each subgroup are more similar to each other than they are to members not in the subgroup has been extensively studied through the statistical method of cluster analysis. Such an analysis has been previously used to group and characterize student responses to written questions about two-dimensional kinematics (Witmann, 2002) as well as multiple-choice tests (Ding and Beichner, 2009). Such authors outline that the power of cluster analysis lies in the clusters arising from the data and possibly uncovering unexpected relationships between student responses.

Cluster analysis can separate students into groups that can be recognized and characterized by common traits in students' answers without any prior knowledge by the researcher of what form those groups would take (unbiased classification). However, it must be taken into account that the resulted groups have to reasonably make sense to a researcher, and the probabilistic nature of the methodology.

A recent paper (Stewart *et al.*, 2012) analyses the evolution of student responses to seven contextually different versions of two Force Concept Inventory questions, by using a model analysis for the state of student ability and a clustering method in characterizing the student distribution answers. The paper shows that the clustering algorithm ( $k$ -means clustering) is an efficacious method of examining the subgroup structure of student understanding and it produces significant subgroup population fractions. The authors conclude that the  $k$ -means algorithm is an effective mechanism for extracting the underlying subgroups in student data and that additional insight may be gained from a carefully analysis of clustering results.

Cluster analysis can be achieved by various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them. Notions of clusters include groups with small distances among the cluster members, dense areas of the data space, intervals or particular statistical distributions. The appropriate clustering algorithm and parameter settings depend on the individual data set and intended use of the results.

Clustering methods can be roughly distinguished as hierarchical and non-hierarchical ones. The first category of algorithms is based on the core idea to build a binary tree of the data that successively merges similar groups of points and by visualizing this tree a useful summary of the data can be provided. Data are consequently connected to form clusters based on their distance. The second category of algorithms partitions the data space into a structure known as a Voronoi diagram (a number of regions including subsets of similar data).

In this paper we will describe the main characteristics of a particular non-hierarchical method, called  $k$ -means clustering and will apply this to student answers of an open-answer questionnaire in order to make evident the typical behaviours of a sample of university students in relation to Boltzmann Factor topic.

## 2. – Theoretical framework of data analysis

Open-ended questionnaires often offer insights or issues not captured in closed questions. However, coding student answers can be harder than coding close-ended, or multiple choice, ones. This is mainly due to the fact that in open-ended questionnaires the researchers have to take into consideration all possible answers to the questions, in contrast to the other cases, where a limited list of possible answers to each item is already provided. Generally, techniques developed for analysing qualitative data are used to analyse and code the responses to open-ended questions. Researchers carefully

TABLE I. – *Data matrix for analysis. The  $n$  students are indicated as  $S_1, S_2, \dots, S_n$ , and the  $m$  answering strategies as  $AS_1, AS_2, \dots, AS_m$ .*

Strategy	Student				
	$S_1$	$S_2$	...	...	$S_n$
$AS_1$	1	...	...	...	...
$AS_2$	1	...	...	...	...
...	0	...	...	...	...
$AS_5$	1	...	...	...	...
...	0	...	...	...	...
$AS_m$	0	...	...	...	...

read responses so to examine patterns and trends and to find common themes emerging from them. Then, these themes, resuming the different trends found in the responses, are developed in a number of categories, that can be considered the typical “answering strategies” put into action by students when fronting the questionnaire items. Therefore, it is possible to include the whole set of answers given to the questionnaire in a limited number of answering strategies, making easier the subsequent coding.

It is often advisable that more than one researcher perform the search for patterns and trends in student answers, and then the whole group of researchers contrast and compare their own findings, in order to reach a consensus of a common table of student answering strategies to be used for the subsequent study.

If  $m$  answering strategies are identified in the total group of answers, each student can be identified by an array,  $a_i$ , composed by  $m$  components 1 and 0, where 1 means that the student used a given answering strategy to respond to an item, and 0 means that he/she did not use it. If we have  $n$  student that answered the questionnaire, a  $m \times n$  binary matrix (the “matrix of answers”) can be built. This matrix is modelled on the one shown in table I. In it, the columns report the  $n$  student arrays,  $a_i$ , and the rows represent the  $m$  components of each array, *i.e.* the  $m$  answering strategies.

For example, let us say that student  $S_1$  used answering strategies  $AS_1, AS_2$  and  $AS_5$  to respond to the questionnaire questions. Therefore,  $S_1$  column in table I will contain the binary digit 1 in the three cells corresponding to these strategies, while all the other cells will be filled with 0.

The matrix depicted in table I contains all the information needed by the researcher to describe the behaviour of students with respect to the subjects dealt with the questionnaire items. However, it needs some elaboration in order to make this information understandable and to classify student behaviour in different groups, or clusters through Cluster Analysis (Ding and Beichner, 2009).

Cluster Analysis (CLA) is defined as the task of grouping a set of elements so that elements in the same group are more alike, in a sense or another, to each other than to elements included in other clusters. Introduced in Psychology by R.C. Tyron in 1939 (Tryon, 1939), CLA has been the object of research interest since the beginning of the 60s of the last century, with a first systematic use due to Sokal e Sneath (Sokal, Henry and Sneath, 1963). Applications of techniques related to CLA are common in many fields,

including informatics, biology, medicine, archaeology Econophysics and market research (Ott, 1999; Allen and Goldstein, 2013; Mantegna, 1999; Cowgill and Harvey, 1999). Whenever it is necessary to classify a large amount of information into distinguishable groups, CLA is an effective and essential method. CLA techniques allow the researcher to locate, within a set of objects of any nature, subsets, or clusters, which have a strong tendency to be homogeneous “in some sense”. The result of the analysis should, in line with the criteria chosen, highlight a high homogeneity within the group (intra-cluster) and high heterogeneity between groups (inter-cluster).

In our case, the student groups are analysed in order to deduce their distinctive characteristics and to find similarities and differences between them. Each cluster is characterized by means of a careful read of the typical trends in answers of the students that are part of the cluster. Other studies (Fazio *et al.*, 2013; Fazio *et al.*, 2012) instead find clusters by comparing each student (*i.e.* the  $m$ -component array containing his/her answering strategies) with researcher-built arrays, representing ideal profiles of student behaviour. These profiles are often known from previous research and the related arrays are characterized by well-defined answering strategies.

**2.1. Distance indexes.** – The clustering procedures need the definition of new quantities that are used to build the clusters, as, for instance, the “similarity” or “distance” indexes. These indexes are defined by starting from the  $m \times n$  binary matrix discussed above.

The similarity between two elements is often expressed in the literature by taking into account the distance,  $D(a_i, a_j)$ , between them (that actually expresses their “dissimilarity”, in the sense that the higher the distance between the elements, the lower is their similarity).

The distance index is often defined by starting from the Pearson’s correlation coefficient,  $R$ . It allows the researcher to study the correlation between two elements,  $i$  and  $j$ , of a set, but the related variables must be numerical.

If we want to deal with two elements identified by non-numerical variables (for example, the arrays  $a_i$  and  $a_j$  containing the binary coding of answers of students  $i$  and  $j$ , respectively), we can use a modified form of  $R$ , defined in terms of the properties of the elements (*i.e.* the numbers of 1’s and 0’s in the array). A possible definition we propose is

$$R_m(a_i, a_j) = \frac{n_p(a_i \cap a_j) - \frac{n_p(a_i)n_p(a_j)}{N_p}}{\sqrt{n_p(a_i)n_p(a_j) \left(\frac{N_p - n_p(a_i)}{N_p}\right) \left(\frac{N_p - n_p(a_j)}{N_p}\right)}}$$

that is known as “modified Pearson’s coefficient” (Tumminello *et al.* 2011), where  $n_p(a_i)$ ,  $n_p(a_j)$  are the number of properties of  $a_i$  and  $a_j$  that we want to take into account, respectively (the numbers of 1’s or 0’s in the arrays  $a_i$  and  $a_j$ , respectively),  $N_p$  is the total number of properties to study (in our case, the  $m$  possible answering strategies) and  $n_p(a_i \cap a_j)$  is the number of properties common to both  $a_i$  and  $a_j$  (the common number of 1’s or 0’s in the arrays  $a_i$  and  $a_j$ ).

The choice of the type of metrics to use for the distance calculations is often complex and depends on many factors. If we want that two elements  $a_i$  and  $b_j$ , negatively correlated, are more dissimilar with respect to two elements positively correlated (as it is often advisable in research in education), a possible definition of the distance between

$a_i$  and  $b_j$ , making use of the modified correlation coefficient,  $R_m(a_i, b_j)$ , is

$$D(a_i, b_j) = \sqrt{2(1 - R_m(a_i, b_j))}.$$

It is a Euclidean metrics (Gower, 1966; Leisch, 2005) as it is needed to represent the clusters in graphical form.

Once a metric is chosen, a distance between two elements equal to zero means that they are completely similar, while a distance equal to 2 shows that the elements are completely dissimilar. It is, then, possible to construct a new matrix, containing all the distances between the elements of the set. It clearly has the main diagonal composed by 0's (the distance between an element and itself is zero) and it is symmetrical with respect to the diagonal.

**2.2. *k-means method.*** – Non-hierarchical clustering is used to generate grouping of a set of elements by partitioning it and producing a smaller set of non-overlapping clusters having no hierarchical relationships between them. Various algorithms can be used to build the clusters. Among the currently used ones we consider the *k-means*, first proposed by MacQueen in 1963 (MacQueen, 1963)

In the *k-means* algorithm, the starting point is the choice of the number of clusters one wants to populate and of an equal number of “seed points”, randomly chosen between the elements of the dataset. The elements are, then, grouped on the basis of the minimum distance between them and the seed points. The part of a given cluster (the elements of a given cluster) is used to find a new point, representing the average position of the spatial distribution of the cluster elements. This is done for each cluster and the resulting points are defined the cluster *centroids*. The process continues by again grouping each set elements on the basis of the minimum distance between them and the cluster centroid  $C_k$  and re-calculating the average positions of the elements of the new clusters (*i.e.* the new cluster centroids). The iteration ends when the new centroids have the same position of the old ones. The spatial distribution of the set elements can be represented in a two-dimensional space, originating what is known as the *k-means* graph.

Each centroid  $C_k$  defines its cluster and can be used to characterize it. Particularly, if we are able to find an array  $\bar{a}_k$ , of the same dimension of the ones associated to the real cluster elements,  $a_i$ , (*i.e.* the  $m$  answering strategies to the questionnaire) and composed by 0 and 1 values, we can consider it as a new cluster element (in our case, a student) summarizing the average characteristics of the real cluster elements. We can, then, study the answering strategies composing the  $\bar{a}_k$  array and give sense to the average behaviour of the cluster elements.

In order to find the  $\bar{a}_k$  arrays components of the centroids  $C_k$ , we elaborated a methodology that consists in repeating *the k-means* procedure in reverse. As it is difficult to do this analytically, we used an iterative method that finds the specific array  $\bar{a}_k$ , starting from a generic array with the same dimension. The array  $\bar{a}_k$  found at the end of the iterative method represents the centroid  $C_k$ .

It is possible to verify that the  $\bar{a}_k$  arrays have 1 values exactly in correspondence to the answering strategies most frequently given by the cluster elements. In fact, since a centroid is defined as the geometric point that minimizes the sum of the distances between it and all the cluster elements, by minimizing this sum the correlation coefficients between the cluster elements and the centroid is maximized (see the Gower metrics definition) and this happens when each centroid has the largest number of common strategies with all the elements part of its cluster.

The problem of *a priori* choosing the initial positions of centroids can be solved by repeating the clustering procedure for several values of the initial conditions and selecting those that lead to the minimum values of the distances between each centroid and the elements part of its cluster. Particularly, we found that, in order to find steady values for these minima it may be necessary to repeat the procedures to find a centroid up to 100000 times (depending on the specific spatial distribution of the cluster elements), each one with different initial conditions.

### 3. – Clustering students' answers about the Boltzmann factor

In this section we will apply the techniques above discussed to the analysis of student answers to the open-ended questionnaire reported in the appendix, where are also reported the answering strategies pointed out at the end of the qualitative analysis described in the following.

**3.1. Context and sample.** – The questionnaire was administered to 118 students of 8 classroom of Scientific Upper Secondary School (18 years old) called in Italy *Liceo*. This specific typology of students was chosen for their good motivation and their high level of Mathematics, Chemistry and Physics competences certified by the evaluation of their teachers.

The questionnaire consists of six-items focused on the ability to create explanation related to the physics context of the Boltzmann factor. Students are required to clarify the physical meaning of the quantities involved in a given phenomenon (the evaporation of a water puddle at different temperatures), discuss the related explicative model(s), and propose other experimental situations that can be explained by using the same model(s). The focus is on systems for which a process is thermally activated by overcoming a well-defined potential barrier  $E$ , and is therefore described by an equation containing the Boltzmann factor  $e^{-E/kT}$ , where  $T$  is the system temperature and  $k$  is the Boltzmann constant.

The questionnaire items are inspired by other questionnaires on the processes of modelling already used in previous research (Fazio *et al.* 2012; Lederman *et al.*, 2002; Fazio and Spagnolo, 2008; Ferri, 2006). Further problems concerning their format and language have been solved by a procedure of face-validation phase (Anastasi, 1988): a group of 9 students from different classrooms (one for each classroom) were asked to preliminarily answer the questions. Then a focus group was conducted with the students, in order to clarify the meaning of their answers and get to the final version of the questionnaire to be used with the research sample.

After the questionnaire was submitted to the 118 students of our sample, each researcher independently read the answers, and wrote down a list of typical answers that, in his/her opinion, the students actually gave to each questionnaire item. The two lists were, then, compared and contrasted in several meetings between the researchers, in order to get to a shared and unique list, reporting the 59 typical answers given by the students to the questionnaire items. The integrated reliability of the analysis was good. Discordances between researcher lists were typically found as a consequence of the different personal disposition of the researchers to synthesize the student answers in a more or less restricted number of typologies. In a few cases discordances were due to different researcher interpretations of student statements.

The complete list of 59 typical answers given by students to the questionnaire items is reported in the appendix.

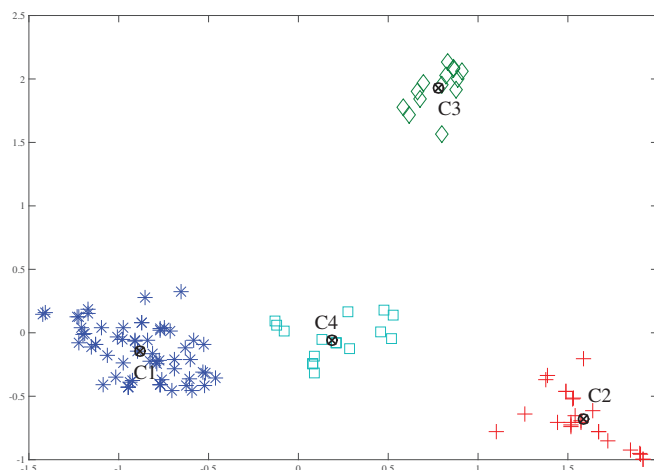


Fig. 1. –  $k$ -means graph. Each point in this Cartesian plane represents a student. Four clusters are clearly depicted and the related centroids ( $C1, C2, C3, C4$ ) are shown.

Each researcher, then, coded the student answers using the previously described list, building a matrix like the one depicted in table I, where  $n = 118$  and  $m = 59$ . Again, some meetings were spent in order to compare the different matrices and to come to a shared one to use for the clustering calculations.

#### 4. – The results

Clustering calculations have been performed by using custom software written in C language, using the  $k$ -means method. The graphical representation of clusters has been obtained by using the well known MATLAB (MATLAB, 2015) software.

Figure 1 shows the representation of the partition of student data set in a 2-dimensional graph, where the  $x$  and  $y$  axis simply report the values needed to place each sample element according with its mutual distance with respect to the other elements.

The clusters obtained by applying the algorithms previously described are characterized by the related centroids, that, as discussed above, are the four points in the graph whose arrays  $\bar{a}_k$  contain the answering strategies most frequently given by the elements of the related clusters. These strategies are defined as follows:  $\bar{a}_1$ : (1F, 2D, 3G, 4C, 5C, 6E),  $\bar{a}_2$ : (1B, 2B, 3A, 4A, 5B, 6A),  $\bar{a}_3$ : (1I, 2G, 3N, 4H, 5G, 6L),  $\bar{a}_4$ : (1D, 2C, 3B, 4A, 5C, 6B), for the centroids  $C_1, C_2, C_3, C_4$  respectively, where the codes in parenthesis refer to the answering strategies to the questionnaire items reported in the appendix.

An analysis of these strategies allows us to characterize the answering strategies common to subjects of the related cluster. In particular, the cluster identified by centroid  $C_3$  is the one composed by the students that exhibit the highest level answering strategies with respect to concepts dealt with the questionnaire. In fact, in average these students correctly recognize the characteristic discussed quantities of physic-chemical phenomena (2G) and are able to identify the relations between even if, in some cases only at a macroscopic level (4H). These students also well know how to find similarities between different systems (5G), providing an explanation of the working mechanism at a microscopic level.



The student characterized by this profile show generalization abilities (5G).

Students of clusters identified by centroids  $C_1$  and  $C_4$  can be defined as intermediate-level ones, even if at slightly different ways. Students in the cluster identified by centroid  $C_1$  are the majority. These students are able to describe the relationships between the quantities at play, preferring the mathematical formulation (2D, 5C, 6E) in some cases only with a macroscopic approach. The explanations of the phenomenon are mostly generic (4C), in fact they never provide a working mechanism of the phenomenon. Some few correct generalizations in different contexts are found just on the basis of mathematical formula (6E).

The student characterized by the centroid  $C_4$  are able to detecting the physical requested quantities using the mathematical formulas (5C) even if, in some cases, the description is referred to everyday life context (4A). Also they fail to determine similarity between apparently different phenomena and therefore are not able to generalize in different contexts (6B).

The student characterized by the centroid  $C_2$  represents the lowest level of involved students. These kinds of students have difficulty in defining and connect each other the physical quantities (2B). Their descriptions of phenomena is made on the basis of their common sense interpretation of phenomena (4A) that, often, is far from the scientific one. They are not able to find possible similarities between phenomena defined in different contexts (6A) and the few tentative of generalization is unfair (5A).

## 5. – Conclusions

In this work we discussed a quantitative method aimed to analyse an open-ended questionnaire on a large sample of students in order to make evident the consistency among student answers. Among the different types of clustering techniques we chose to investigate the non-hierarchical one and in particular the *k-means* method. This technique, not well spread and known in the Education research field, has been briefly described from its theoretical foundations.

We spent particular attention to the definition of the coefficient of correlation and to the type of metric to be used for using the potentiality of this method. The most important result of this work is the definition of a method by which to characterize in a simple way the students behaviour, dividing them into homogeneous groups with the characterization of the centroid obtained through, the *k-means*.

We, moreover, discussed an application of the discussed method to student answers of a questionnaire on a well-known epistemological physical context as the Boltzmann factor. The relevance of using a quantitative method in the Education field emerges from the obtained results. The method permitted us to clearly highlight the student's behaviour and to classify them according to four groups (that emerged *a posteriori* from the data), including students with a similar profiles.

Briefly, the analysis made evident that only a small number of student shows significant ability of generalization. Only few of them are able to built a microscopic model for the description of physical phenomena. The majority of student instead simply provides descriptions more or less related to qualitative analysed phenomena. A limited number of students is bind to common sense explanations. These results are in accordance with the literature related to the generalization and proving abilities of Upper secondary School students in Mathematics and Science (Fazio, C. and Spagnolo, F 2008; Heinze *et al.* 2009; Maaß K 2006; Mariotti, 2006).



The strength of this work is in the possibility to characterize a large students sample through a procedure that does not need any *a priori* assumptions about their typical behaviour. In this sense our approach is different from the typical method quoted in literature (Brousseau, 1997) and it allows researcher to minimize the inevitable subjectivity in the characterization of students behaviour. However it is noteworthy to recall that data quantitatively analyzed are the results of a categorization of raw data (the individual student answers) and such first data reduction can be subjected to errors that, obviously, influences the final evaluation and the inference about the reasoning strategies supporting students' answers. Such errors can be only reduced (through a clear process of coding and successive categorization) and not eliminated and this must be taken into account when we try to infer typical students' reasoning strategies.

## Appendix

Questionnaire items and related answering strategies.

- 1) *A puddle dries more slowly at 20°C than at 40°C. Assuming all other conditions (except temperature) equal in the two cases, explain the phenomenon, pointing out what the fundamental quantities are for the description of the phenomenon and for the construction of an interpretative model of the phenomenon itself.*
  - 1A The relevant quantities are not identified.
  - 1B The relevant quantities are not identified, but a description/explanation based on common sense is given.
  - 1C The relevant quantities are identified, but they are not used properly to give an explanation.
  - 1D Only temperature is identified as relevant, but the phenomenon is not correctly described.
  - 1E Only temperature is identified as relevant. It is used to give a rough description of the phenomenon.
  - 1F The phenomenon is described by means of the macroscopic variables pressure and volume, but a microscopic model is not identified.
  - 1G The phenomenon is described by means of the macroscopic variables temperature, energy and heat, but a microscopic model is not identified.
  - 1H The phenomenon is described by means of a mathematical formula, but a microscopic model is not identified.
  - 1I The phenomenon is not adequately described (by means of a mathematical formula or verbally), but a microscopic "functioning mechanism" is roughly presented in terms of "molecular collisions".
  - 1L The phenomenon is not adequately described (by means of a mathematical formula or verbally), but a microscopic "functioning mechanism" is presented in terms of energy exchange between molecules.
  - 1M The phenomenon is verbally described and a microscopic "functioning mechanism" is roughly sketched.
  - 1N The phenomenon is described by means of mathematical relations between macroscopic quantities and a microscopic "functioning mechanism" is found.

- 2) *In chemical kinetics it is well known that the rate of a reaction,  $u$ , between two reactants follows the Arrhenius law:*

$$u = Ae^{-\frac{E}{kT}}.$$

*Describe each listed quantity, clarifying its physical meaning and the relations with the other quantities.*

- 2A The fundamental quantities are not described and/or only examples of its application to everyday-life phenomenology are given.
- 2B Some quantities are mentioned, but no description of the process is given.
- 2C The relevant quantities are found, but only a few are described in terms of their physical meaning.
- 2D The relevant quantities are found, but only described in terms of their mathematical meaning in the formula. No relation between them is identified.
- 2E The relevant quantities are found and correctly described in terms of their physical meaning. No relation between them is identified.
- 2F The relevant quantities are found and correctly described in terms of their physical meaning. Some relations between them are identified.
- 2G The relevant quantities are found and correctly described in terms of their physical meaning. The relations between them are correctly identified.
- 3) *What do you think the role of a catalyst is, in the development of a chemical reaction?*
- 3A A definition of catalyst is given, which does not conform to the scientifically correct one.
- 3B A definition of catalyst based on an analogy with the concept of enzyme is given. The analogy is recalled without providing additional reasoning.
- 3C The catalyst is described as a substance which speeds up a chemical reaction. No additional explanation is supplied.
- 3D The catalyst is described as a substance which shifts the chemical equilibrium towards the products. No additional explanation is supplied.
- 3E The catalyst is described as a substance which speeds up a chemical reaction. An explanation is given using common language.
- 3F The catalyst is presented as a substance which shifts the chemical equilibrium towards the products. An explanation is given using common language.
- 3G The catalyst is presented as a substance which speeds up a chemical reaction. The concept is generically described in terms of energy.
- 3H The catalyst is presented as a substance which shifts the chemical equilibrium towards the products. The concept is generically described in terms of energy.
- 3I The catalyst is presented as a substance which speeds up a chemical reaction. The concept is described by simply citing the energy gap concept, without any explanation.

- 3L The catalyst is presented as a substance which shifts the chemical equilibrium towards the products. The concept is described by simply citing the energy gap concept, without any explanation.
- 3M The role of a catalyst in a chemical reaction is discussed referring to the energy gap concept, but only in macroscopic terms.
- 3N The role of a catalyst in a chemical reaction is discussed taking into account the energy gap concept. The concept is explained considering a microscopic model regarding collisions between molecules.
- 3O The role of a catalyst in a chemical reaction is discussed taking into account the energy gap concept. The concept is explained considering a microscopic model which links the energy gap concept with the molecular energy.
- 4) *Can you give your own physics model of the Arrhenius law?*
- 4A Everyday-life concepts are mentioned, without any correct relation to the Arrhenius law.
- 4B Scientific concepts, such as energy, temperature or molecular thermal agitation, are mentioned, but they are not correctly related to the Arrhenius law.
- 4C Arrhenius law is described as a mathematical function of  $T$  or  $E$ . No explanation of the meaning of these quantities is given.
- 4D Arrhenius law is described as a mathematical function of both  $T$  and  $E$ . No explanation of the meaning of these quantities is given.
- 4E Arrhenius law is described as a function of both  $T$  and  $E$  and the meaning of these two quantities is outlined mainly in mathematical terms.
- 4F Arrhenius law is described as a function of both  $T$  and  $E$ . The physical meaning of these two quantities and/or of their ratio in the Arrhenius law is outlined.
- 4G Arrhenius law is described outlining the physical quantities involved. Collision theory is sometimes mentioned, but a clear reference to a microscopic model is not always present.
- 4H A generic explanation based on a microscopic model of collisions between molecules is given. The activation energy concept is outlined but its relation with  $kT$  is not clearly presented.
- 4I A quantitative explanation in terms of the "collision theory" is given. A correct microscopic model is presented and the role of the activation energy and of  $kT$  is clearly expressed.
- 5) *Can you think of other natural phenomena which can be explained by a similar model?*
- 5A A few phenomena not related to the model are mentioned. No explanation is given.
- 5B A few phenomena not related to the model are mentioned. An explanation is given using common language.
- 5C A few phenomena not related to the model are mentioned. An explanation is given using mathematical formulas.

- 5D Some phenomena related to the model are mentioned, and non-chemical phenomena are also taken into account, but a clear explanation is not given.
- 5E Some phenomena related to the model are mentioned, and non-chemical phenomena are also taken into account. An explanation is given using mathematical formulas.
- 5F Some phenomena related to the model are mentioned, and non-chemical phenomena are also taken into account. An explanation is given outlining a common microscopic model, but energy and temperature are not clearly interrelated.
- 5G Some phenomena related to the model are mentioned, and non-chemical phenomena are also taken into account. An explanation is given outlining a common microscopic model. The role of energy and temperature in the model is clearly discussed.
- 6) *Which similarities can be identified in the previous phenomena? Is it possible to find a common physical quantity which characterizes all the systems you discussed in the previous questions?*
- 6A No similarities are detected and questions 1) and 2) are identified as being related to a different context on the basis of everyday-life reasoning.
- 6B No similarities are detected and questions 1) and 2) are identified as being related to a different context. An explanation is given, mentioning physical quantities which are not really relevant to the correct explanation of the questions.
- 6C A few correct similarities are found, but physical quantities are given, which are not really relevant to the correct explanation of the questions.
- 6D Incorrect similarities are found on the basis of a mathematical formula.
- 6E A few correct similarities are found on the basis of a mathematical formula.
- 6F Correct similarities are found, but  $E$  and  $T$  are not always considered common to all phenomena.
- 6G Some correct similarities are found.  $E$  or  $T$  is considered to be characteristic of the various phenomena, but a clear justification is not given.
- 6H Some correct similarities are found.  $E$  or  $T$  is considered to be characteristic of the various phenomena, clearly explaining why.
- 6I Some correct similarities are found.  $E$  or  $T$  is considered to be characteristic of the various phenomena, but the relevance of their ratio in explaining the energy threshold processes is not clearly presented.
- 6L Some correct similarities are found.  $E$  or  $T$  is considered to be characteristic of the various phenomena. The activation energy role is correctly discussed in all the mentioned phenomena, but only in macroscopic terms.
- 6M Some correct similarities are found.  $E$  or  $T$  is considered to be characteristic of the various phenomena. The activation energy role is correctly discussed in all the mentioned phenomena, on the basis of a microscopic model.

## REFERENCES

- [1] ALLEN D. N. and GOLDSTEIN G. (Editors), *Cluster Analysis in Neuropsychological Research: 13 Recent Applications*, (Springer Science+Business Media, New York) 2013, DOI 10.1007/978-1-4614-6744-1.2.
- [2] ANASTASI A., *Psychological testing* (McMillan, New York) 1988, p. 144.
- [3] BAO L. and REDISH E. F., *Phys. Rev. ST Phys. Educ. Res.*, **2** (2006) 010103.
- [4] BAO L., HOGG K. and ZOLLMAN D., *Am. J. Phys.*, **70** (2002) 766.
- [5] BROUSSEAU G., *Theory of Didactical Situations in Mathematics 1970-1990/by Brousseau*; edited and translated by BALACHEFF N., COOPER M., SUTHERLAND R. and WARFIELD V. (Springer Netherlands) 1997.
- [6] COWGILL M. C. and HARVEY R. J., *Comput. Math. Appl.*, **37** (1999) 99.
- [7] DING L. and BEICHNER R., *Phys. Rev. ST Phys. Educ. Res.*, **5** (2009) 020103.
- [8] FAZIO C. and SPAGNOLO F., *S. Afr. J. Educ.*, **28** (2008) 469.
- [9] FAZIO C., DI PAOLA B. and GUASTELLA I., *Phys. Rev. ST Phys. Educ. Res.*, **8** (2012) 010110.
- [10] FAZIO C., BATTAGLIA O. R. and DI PAOLA B., *Phys. Rev. ST Phys. Educ. Res.*, **9** (2013) 020101.
- [11] FERRI R. B., *ZDM*, **38** (2006) 86.
- [12] GOWER J. C., *Biometrika Trust*, **53** (1966) 325.
- [13] HEINZE A. and REISS K., *Teaching and learning proof across the grades: AK-16 Perspective*, edited by STYLIANOU D. A., BLANTON M. L. and KNUTH E. J. (Routledge, New York) 2009, pp. 191–203.
- [14] LEDERMAN F., ABD-EL-KHALICK R. L., BELL and SCHWARTZ R., *J. Res. Sci. Teach.*, **39** (2002) 497.
- [15] LEISCH F., *Comput. Stat. Data Anal.*, **51** (2006) 526.
- [16] MAAß K., *ZDM*, **38** (2006) 113.
- [17] MACQUEEN J., in *Proc. 5th Berkely Symp. Math. Statist. Probab. 1965/66*, edited by LECAM L.M. and NEYMAN J., Vol. **I** (University of California Press, Berkely) 1967, pp. 281–297.
- [18] MANTEGNA R. M., *Eur. Phys. J. B*, **11** (1999) 193.
- [19] MARIOTTI M. A., in *Handbook of research on the psychology of mathematics education: Past, present and future*, edited by GUTIÉRREZ ANGEL and BOERO PAOLO (Sense Publishers, Rotterdam/Taipei) 2006, pp. 173–204.
- [20] MATLAB version 8.5 Natick, Massachusetts: The MathWorks Inc., 2015.
- [21] OTT J., *Analysis of Human Genetic Linkage*, third edition (Johns Hopkins University Press, Baltimore, London) 1999.
- [22] SOKAL R. R., HENRY P. and SNEATH A., *Principles of Numerical Taxonomy* (W.H. Freeman, San Francisco) 1963.
- [23] STEWART J., MILLER M., AUDO C. and STEWART G., *Phys. Rev. ST Phys. Educ. Res.*, **8** (2012) 020112.
- [24] TUMMINELLO M., MICCICHÉ S., DOMINGUEZ L. J., LAMURA G., MELCHIORRE M. G., BARBAGALLO M. and MANTEGNA R. N., *PLoS ONE*, **6** (2011) e23377, doi: 10.1371/journal.pone.0023377.
- [25] TRYON R. C., *Cluster Analysis: Correlation Profile and Orthometric (factor) Analysis for the Isolation of Unities in Mind and Personality*, edited by ARBOR ANN (Edwards Brothers) 1939.
- [26] WITTMANN M. C. and SCHERR R. E., in *Physics Education Research Conference Proceedings 2002*, edited by CUMMINGS K. C., MARX J. and FRANKLIN S. (PERC Publishing, Boise, ID, USA) 2002, p. 43–46.