Modeling of data - An interdisciplinary approach

Mun Fie Tsoi, Ngoh Khang Goh and Lian Sai Chia

Science and Technology Education, National Institute of Education,
Nanyang Technological University, Republic of Singapore
E-mail: mfrtsoi@nie.edu.sg

Abstract

In the context of symbolic representation of data, learning to connect the data generated from macroscopic events with hypothesized changes taking place at the atomic level and the symbolic level used to describe these events is a difficult task for many students. Therefore, an interdisciplinary approach to modeling of data is one way to address this problem. This paper provides insights on modeling of data using DataStudio, a data logging and analysis tool. It concentrates on the overall instructional design of the interdisciplinary approach in the field of mathematics and chemistry that maximize the potential of DataStudio to enhance subject content and pedagogy. Modeling of physicochemical phenomena will be illustrated with examples. Implications for designing this interdisciplinary approach to manipulation and transformation of data will also be discussed.

Introduction

In science and chemical education, three understanding levels have been identified (Johnstone, 1991; Gabel, 1993). They are (a) the macroscopic level, which deals with sensory/ visible phenomena such as laboratory observations and data; (b) the microscopic level, which deals with particles such as atoms, ions and molecules; and (c) the symbolic level, which represents the matter in terms of chemical formulae and equations. Ben-Zvi & Gai (1994) have also suggested that teaching chemical concepts in the particulate nature of matter should emphasize the relations between these concepts and real world phenomena on both the macro and micro level. In some respects, these three understanding levels (macroscopic, microscopic, and symbolic) are also similar to the three modes of repesentation of mathematical knowledge identified in mathematics and mathematical education (Bruner, 1966). They are (a) ikonic level (imagery), which deals with pictorial representations; (b) enactive level (action), which deals with actual events; and (c) symbolic level (language), which represents the operations or mathematical models in terms of symbols and equations.

For both chemical and mathematical education, students are required to demonstrate transfer between the phenomenon, macro or iconic and its atomic or enactive events and symbolic representations. However, students do find difficulties in establishing such connections. As such, an interdisciplinary approach to modeling of data using DataStudio, a data logging and analysis tool by *Pasco Scientific* would address this problem. Besides, this approach would also provide opportunities for the creation of dynamic, symbolic representations of nonconcrete, formal constructs that are frequently missing in the mental models of these students.

With this in mind, an instructional design based on constructivist paradigm is then developed focussing on modeling of data via an interdisciplinary approach. This seems compatible with the constructivist paradigm in that it establishes a situation in which students have ample opportunities to interact with an information technology rich environment and make meaningful connections in their knowledge schemata. This constructivist paradigm for instructional design focussing on modeling of data is illustrated in Table 1.

Table 1. Constructivist paradigm for instructional design

Instructional Problem: To learn how to model data	
	Constructivist Paradigm
Goals	Each learner specifies what their individual goal is in relation to this problem
Assessment	Instructor, self evaluation through anecdotal or laboratory reports of data collected and analysed
Strategies	Exploration, interpretation and making decisions about authentic tasks
Delivery Systems	An information technology learning environment providing relevant, interesting and real-world exploratory hands-on learning on the modeling of data

As such, certain aspects of ISD (Instructional Systems Design) would have to become more flexible when a constructivist paradigm is adopted (Bednar et al., 1995). In this context of modeling of data, content analysis is not important because content cannot be prespecified. Domains can be defined, but specific objectives must come from the students' perception of relevancy. In other words, the generic ISD model also known as the AADIE Model that involves an integrated set of steps (analysis, design, development, implementation, and evaluation) evolved from general systems theory would have to be adapted to this paradigm. Development of multiple perspectives on a learning task should be encouraged including an awareness of the interdisciplinary aspects from both the chemical and mathematical disciplines. As the focus is on the learner's level of reflectivity and individuality, an analysis of representative learners is not necessary.

Modeling of data

In these aspects, DataStudio would have the potential to assist learners in learning how to learn in the realm of chemical and mathematical domains where scientific data is collected and analysed in real time. DataStudio provides the tools for modeling data. The main screen of DataStudio is shown in Figure 1. The major components of the main screen are labelled. Data could be modeled in two ways using the functional tools, *Calculate* function and *Curve Fit* function. The *Calculate* function is used to construct a model of an equation of a physical phenomena that would help one understand its mathematical and physical nature whilst the *Curve Fit* function is used to construct a model to find a suitable mathematical expression to describe one's data, and thus to better understand the nature or the system one have just studied in the laboratory.

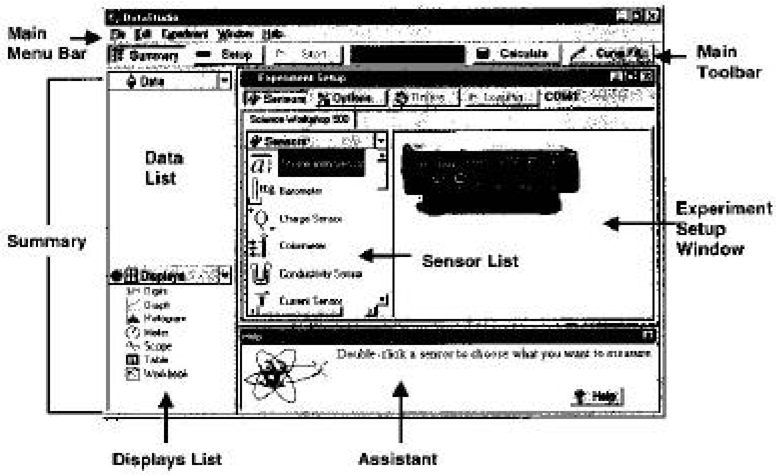

Figure 1. DataStudio main screen

In the experimental activity of Boyle's Law in which the pressure of a gas in a container is related to the volume of the gas, one could elicit higher-order responses using thinking skill such as prediction. For example, "What happens to the pressure in a container of air as its volume is changed while the temperature remains constant?" could be asked. For this activity, students would be able to (a) use a Pressure Sensor to measure the change in pressure and of the air inside a syringe as the piston in the syringe causes the volume of the air to change; (b) use the software to enter values of volume and record the corresponding pressures; and (c) interpret the data to determine the relationship of pressure and volume for a gas at a relatively constant temperature.

Real-time data collected would then be manipulated using the *Calculate* function and *Curve Fit* function. The *Calculate* function is best used for manipulating and transforming data. Using the various function menus of Scientific, Statistical, and Special, one could create new calculations based on the data. Concurrently, one could also use the *Curve Fit* function to fit the data to a curve. A variety of curves could be fitted to the data set. Once the curve fit expression is selected from a pool of 15 pre-defined equations, the software determines the fitting parameters using the method of least-squares.

The sample data modeled is illustrated in Figure 2. The top graph shows the plot of volume versus pressure whilst the bottom graph shows the plot of inverse volume versus pressure. From the chemical discipline perspective, kinetic theory could be applied. Concurrently, from the mathematical discipline perspective, imagery could be promoted by the use of these well-organised graphical representations. Such data could also be further modeled by plotting the graph of pressure, P times volume, V versus pressure, P in order to consider a primary source of error which is the volume of air, A unaccounted for in the tubing and pressure chamber of the sensor. This is illustrated as follows:

$$P(V+A) = k$$
$$PV + AP = k$$
$$PV = k - AP$$

Thus, the gradient of this graph would give the value of this primary source of error, A. So, by creating a new quantity called total volume, Vtot =V + A, a reasonable value for A, say A = 0.5 ml could be added to the volume of air in the syringe as a correction factor. In fact, as a form of verification, the plot of P x Vtot versus P would give a gradient of zero i.e. a horizontal line would be obtained.

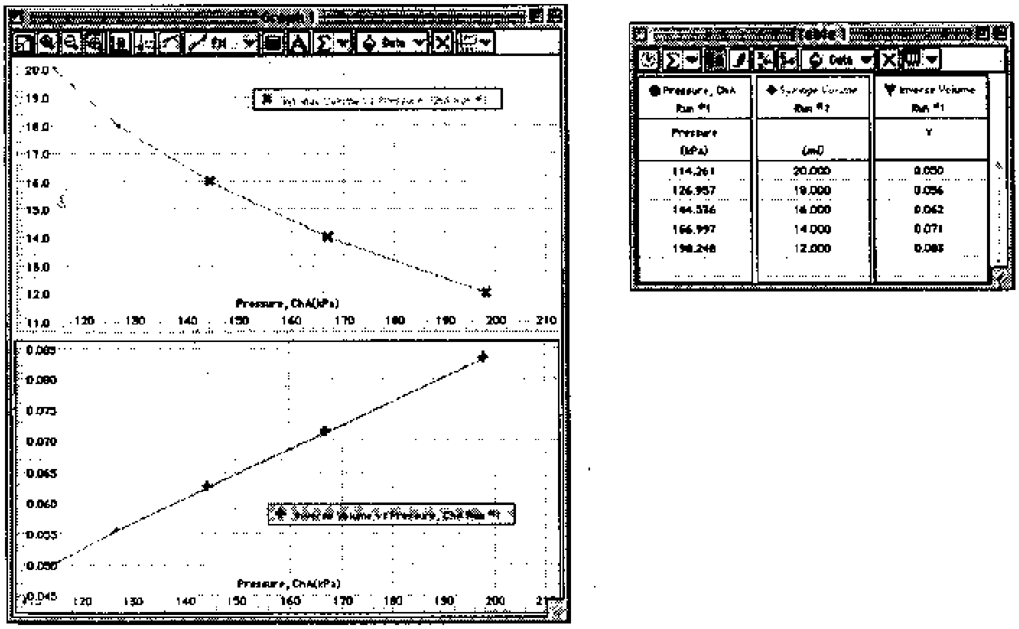

Figure 2. DataStudio sample data (Boyle's Law)

Besides, Boyle's Law is also verified both at the microscopic level of kinetic theory and the ikonic level of graphical representations through this interdisciplinary approach to modeling of real-time scientific data.

As such, these levels of understanding could also be extended to the activity of conductivity in which ionic and nonionic molecules are explored. The objective is to compare the conductance of (a) an ionic molecule that has one equivalent of chloride when dissolved in water; (b) an ionic molecule that has two equivalents of chloride when dissolved in water; and (c) a molecule that is not ionic when dissolved in water. Similarly, higher-order thinking question on the relationship between the conductivity of a solution and the number of negative ions in solution could be established.

The experimental data obtained for the solutions of sucrose, sodium chloride and calcium chloride in terms of conductivity is illustrated in Figure 3. Measuring the conductance of a dilute solution of known molarity could give a good insight into the properties of the dissolved molecule. These experimental observations could also aid the understanding of chemical concepts such as molar versus normal solutions and equivalent weights of a solute.

The concept that molecules vary as to their properties when they are dissolved in an ionic solution would be much appreciated. For instance, the ions of sodium chloride when dissolved in water move about freely in solution making the solution conductive of electricity. Molecules of calcium chloride dissolve to yield more than one chloride anion resulting in more conductance per mole of non-dissolved molecule, compared with the conductance of a mole of a molecule that dissolves to form only one anion per molecule. Some molecules such as sucrose do not dissociate into ions when they are dissolved in water. These solutions have very low conductance.

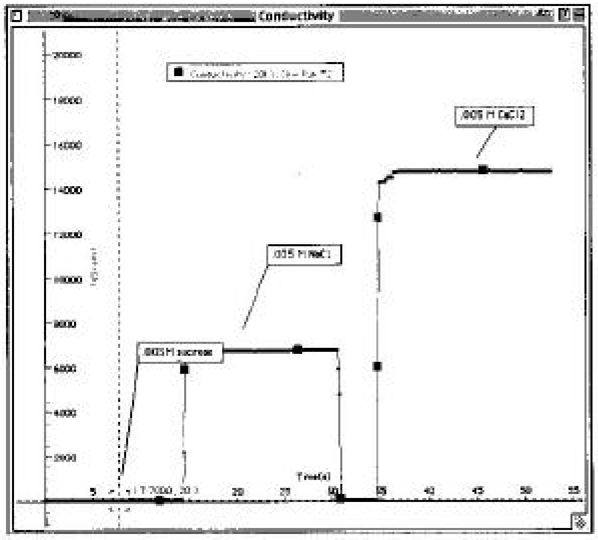

Figure 3. DataStudio experimental data (Conductivity)

In this instance, for both activities students are to infer the characteristics of the model underlying the physicochemical phenomena. This is done using the interdisciplinary approach to modeling of real-time scientific data.

Implications

Instructional Systems Design (ISD) has in a way been affected by theories of learning especially from the cognitive science. However, there has also been an interest in building instruction to facilitate not only the thinking processes but also the interdisciplinary processes. Therefore, a constructivist paradigm for instructional design may be appropriate for modeling of scientific data using the interdisciplinary approach. This paradigm claims that learning is more than conditioning or acquired knowledge, rather it is constructed knowledge. This means that learners interpret information in the context of their own experiences. Learning should be personalised, set in authentic contexts, and oriented to problem solving.

In this case, the constructivist paradigm for the instructional design would consist of the following components: (a) learning is personal discovery based on insight; (b) type of learning is analysing or problem solving; (c) instructional strategies are provided for active and reflective learner; (d) media strategy is a responsive information technology learning environment; and (e) the key concept is autotelic principle (intrinsic motivation). Therefore, an interdisciplinary approach to modeling of data using the DataStudio, a data logging and analysis tool would address this learning situation to a large extent.

For effective modeling of scientific real-time data, it is essential that not only the procedural knowledge but also the declarative knowledge need to be transmitted explicitly in a logical and systematic sequence. The entire learning process has to be explicit and constructive. Students should be explicitly taught these domains of knowledge, emphasising the three levels of understanding identified in chemical and mathematical education through which they could construct mental models that, in turn would enable them to understand the presented conceptual models.

For example, when student observes certain trends in experimental data, he/she could fit data to a curve using the *Curve Fit* function to develop models that could lead to general conclusions about physicochemical phenomena. In terms of instructional pedagogy, students need to be reflective and self-regulating by asking themselves, "what does the experimental data look like?" and "does it look like a parabola, a straight line, a sine curve, an exponential, or a combination of functions?" Besides, students need to understand that it is wise to choose the most general form of function when fitting data to a curve and that the best fit to the data must also make physical sense.

In other words, the interdisciplinary approach to the modeling of data has to be understood as a holistic integration of the three levels of understanding identified in chemical and mathematical disciplines as macroscopic or ikonic; microscopic or enactive; and symbolic respectively. This approach is similar in some aspects to that of the modeling process as the 'learning of a series of steps to identify only those essential features of the system, and to evaluate, according to specific rules, the selected model' (Halloun, 1996) or even that of an integrated reasoning process in which visual and analogical modeling are used in the creation of informal representations of a problem (Nersessian, 1995).

Since mental models are personal constructions, it may seem that the most appropriate methodological path for meaningful learning would be to apply this interdisciplinary approach to the modeling of scientific real-time data. This approach would assist students to construct relevant internal representations or mental models consistent with knowledge that is scientifically acceptable and coherent with the conceptual models presented in the laboratory. Such construction of mental models would also enable the students to further understand the physicochemical phenomena according to scientific theories.

Therefore, careful consideration of the following pertinent issues, for example, (a) what kind of cognitive level the learners will be at initially; (b) how best the learning materials or tasks will be organised and structured; (c) what type of basic tools for generation of mental models will be used; (d) what type of process skills will be employed; and (e) how best the thinking framework is to be infused in the instructional pedagogy would need to be fully addressed to realise the potential of this interdisciplinary approach to the modeling of scientific real-time data.

References

Bednar, A.K., Cunningham, D., Duffy, T. M., & Perry, J. D. (1995). Theory into practice: How do we link. In G.J. Anglin (Ed.), *Instructional technology: Past, present and future* (pp. 100-112). Eaglewood, CO: Libraries Unlimited.

Ben-Zvi, N., & Gai, R. (1994). Macro- and micro-chemical comprehension of real-world phenomena. *Journal of Chemical Education*, 71(9), 730-732.

Bruner, J. S. (1966). *Towards a theory of instruction*. Cambridge, Massachusetts : Harvard, University Press.

Gabel, D. L. (1993). Use of the particle nature of matter in developing conceptual understanding. *Journal of Chemical Education*, 70, 193-194.

Halloun, I. (1996). Schematic modelling for meaningful learning of physics. *Journal of Research in Science Teaching*, 33, 1019-1041.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer assisted Learning*, 7, 75-83.

Nersessian, N. (1995). Should physicists preach what they practice? *Science and education*, 4, 203-226.