WHEN DIDACTICS MEETS DATA SCIENCE

Franck Salles, Reinaldo Dos Santos, Saskia Keskpaik

DEPP, Ministry of education, France, franck.salles@education.gouv.fr

The Department for Evaluation, Prospective and Performance (DEPP) at the French ministry of education provides deciders with data analysis on the state of the French educational system. French students' performance is measured through Large Scale Assessments (LSA). From 2018, DEPP's LSA instruments are computer based when it comes to assess secondary school students. In mathematics, new test instruments have been developed embedding digital tools in support of the mathematical task. This paper presents the results of a study aiming at profiling student's mathematical activity, based on a "big-data" analysis of the digital traces left in technology enriched interactive items by a sample of 9-graders in France. Results give hard evidence of how data science in education can benefit from the mixed contribution of three paramount aspects: technology, analytics and didactics.

Keywords: digital, assessment, big data, didactics

CONTEXT AND RESEARCH QUESTIONS

CEDRE (Subject related sample based assessment cycle) is a sample based large scale assessment aiming at measuring students' abilities in Mathematics at the end of Grade 9 every 5 or 6 years in France. Constructed and designed by the Department for Evaluation, Prospective and Performance (DEPP) at the French ministry of education, its framework is based on the national French curriculum in mathematics. First administered in 2008 and 2014, CEDRE will be administered again in May 2019. This new cycle will be computer based for the first time. As trends must be secured so comparison with previous cycles is guaranteed, a large part of test instruments are similar to formerly paper based items. However, the DEPP developed interactive items, very different from more classical item formats, in order to fully profit from the potentialities of assessing with digital tools (Stacey, 2013). Offering students the possibility to use digital tools during the assessment may outsource (Drijvers, 2018) basic procedural work to such tools. Therefore opportunities are given to students to better engage higher order skills such as devising a strategy or mathematical thinking. To what extend can logdata analysis better inform on students' performance in LSA and explain achievement? Can it participate to categorize students' productions and procedures, allowing didactical interpretation and profiling? Log data files issued from this pilot test have been analysed using a combination of machine learning methods enriched by the results of an *a priori* didactical analysis of the mathematical task in each item.

A PRIORI DIDACTICAL ANALYSIS

A conceptual framework was designed in order to *a priori* analyse the mathematical task in interactive items. This analysis appeared necessary from a methodological point of view in order to draw hypothesis and allow defining variables of interest with respect to the potential mathematical activity in the items. The didactical frame was structured around three main questions: How does mathematical knowledge need to be adapted in order to answer the question? What tools' utilization is necessary to solve the problem? How do student/machine interactions influence the mathematical task?

Determining what mathematical knowledge are involved in items is a preliminary necessity to task analysis. Beyond listing them, we have to identify and describe the way they must or could be operated and what operation's adaptions are necessary to achieve the tasks (Robert (2005) et Roditi-Salles (2015)). This level of analysis can be a first step towards determining choices students have to make, the number of step required, types of errors etc.

These very specific types of tasks are set in a technological environment and embark digital tools. Taking into account digital technology leads us to address two complementary aspects: digital tools utilizations and human/machine interactions.

Tool use can be described in reference to the instrumental approach (Rabardel (1995)). This approach distinguishes a tool from an instrument. The tool gets the status of an instrument when it is used by the student as a mean to solve the problem. In this case, its use is decomposed into "*cognitive schemes containing conceptual understanding and techniques for using a tool for a specific type of task*" (Drijvers (2012)). The analysis will then focus on identifying and describing utilization schemes potentially involved in the task. Rabardel distinguishes two types of utilization schemes: "*usage schemes, related to "secondary tasks" (...) and instrument-mediated action schemes,(...) related to "primary tasks" (...) [which] incorporate usage schemes as constituents.*" (Ibid, p 83)

Interactions and feedback are important features of problem solving situations, especially in an assessment context. Being immerged in a digital environment gives birth to specific kind of human/machine interactions. Most of them are intentional and planned by developers when designing the assessment environment, others are not, but all somehow carry information students can grasp to proceed into the solving process. Laborde (2018) distinguishes two types of feedback with regards to technologically enriched situations: one issued from task specific digital tools, another being the "teacher's voice". This last type of feedback is meant to help students catch, adapt and retain information given in the environment. In a summative assessment context it should be very limited as it could interfere with the objective of measuring students' ability. Nonetheless it can be considered paramount in a formative orchestration. If a summative assessment platform as the one used at the DEPP can be considered as a *non-didactical* environment, where the teacher's voice is supposed to be absent, we yet consider this environment as potentially allowing machine/student interactions. In his Theory of didactical situations, Brousseau (1998) separates three levels of feedback depending on the nature of the environment's mathematical reaction: the feedback can either reflect on actions, formulations or validations. This last model has been used to help describing item/student interactions in DEPP's interactive items.

APPLYING THE FRAME TO A SPECIFIC TASK ANALYSIS: "TREE GROWTH".

In the exercise used in the study, two nonlinear functions model two different tree growths. Both are given in linked numerical and graphical representations (Stacey (2013)). Students act in the numerical representation (a table of values) entering the age of the trees in months. A calculation tool returns the corresponding tree heights and a graphing tool spots the points in the graph. Both actions are realized when student press the "calculate and graph" button. By default, values for 300, 500 and 600 months are given. Additional tools can be used: a pencil, an eraser (not allowing erasing given information), a length tool, a compass, a non-scientific calculator. The question can be translated as: "At what age (other than 0 month) do both trees have the same height?"

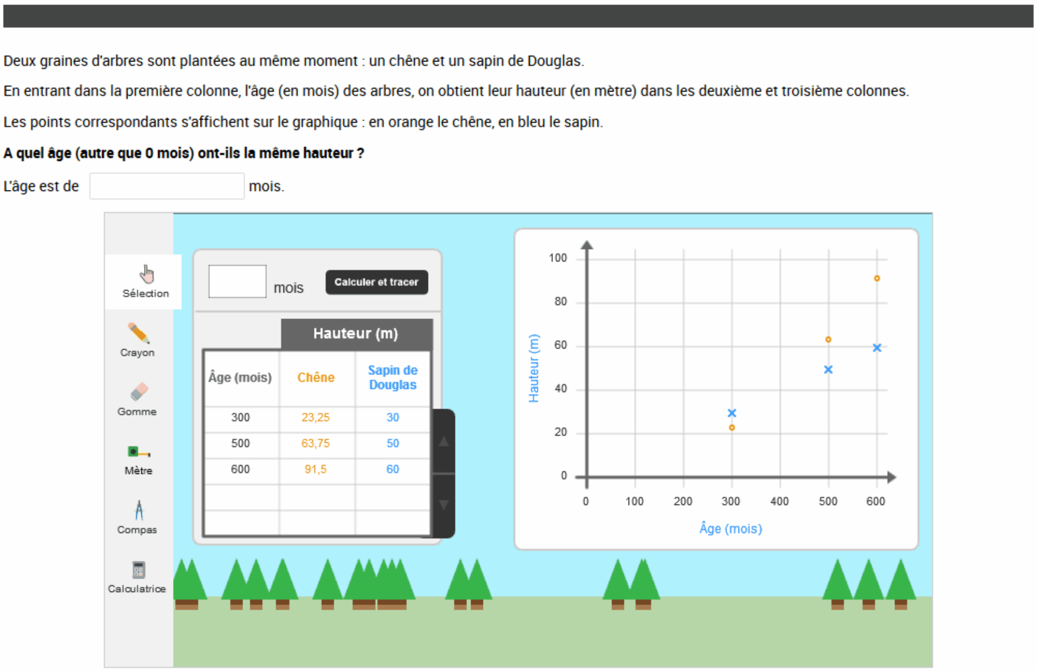

Figure 1: Interactive item "tree growth"

From Grade 9 students' point of view, this task requires conceptual understanding (Kilpatrick and al. (2001)) of functions and their representations (table of values, graph). Functions have two different characters as Sfard (1991) or Drijvers (2012) showed: "*In lower secondary grades, functions mainly have an operational character and are seen as an input–output 'machine' that process input values into output values. In higher grades, functions have a more structural character with various properties (Sfard, 1991). They become mathematical objects that are represented in different ways, are ordered into different types according to their properties, and are submitted to higher-order processes such as differentiation and integration. We argue that the transition from functions as calculation operations to functions as objects is fundamental for conceptual understanding in this domain.*" (Drijvers (2012)). Students can adapt the problem adding intermediary information using the calculation and graphing tool. On one hand, they can then opt for a trial and error method. This would consist in entering a number of months, comparing the results returned either in the numerical or graphical representation, deciding to enter another number of months until the solution (390) is found. Alternating tries around the target value or aiming at it from below or above could improve the trial and error process. These students essentially show good understanding of the concept of functions in their operational character. This method could imply a relatively large number of tries. On the other hand, students having a good understanding of both functions variations, notably from studying them in the graphical representation, can quickly aim at the target number of months. The pencil can for example be used to draw lines and introduce a continuous representation of the functions. The inversion of tree heights between 300 and 500 months can also be noticed. These students comprehend functions as objects with properties, in their structural character.

The following digital tools are at students' disposal within the item:

- A keyboard (with or without number pad) and mouse.
- A "calculation and graph" tool: specific to the item.

- A pencil (common to any item on the platform). Usage: click the starting point, move the mouse to trace, click to stop writing.
- An eraser only allowing to erase pencil traces or measurement tool traces. Usage: clicking erases all pencil traces together.
- A compass
- A calculator

The “calculation and graph tool” does not require complex usage schemes. Two usage schemes are identified for this tool: enter a number of months within the domain [0 ; 600] via an input box and a popup number pad and understand that the tool returns unique heights (outputs) for both trees (numerical and graphical representations) when the button “calculate and graph” is clicked. No tutorial or tool training is proposed to students. Usage schemes are pretty close to relatively usual tools such as a currency converter. Nevertheless, we can imagine some students feeling the need to appropriate the tool by first using it for testing purposes, for example entering extreme values or values not directly connected to the primary task. Using this tool is compulsory to succeed the item. In reference to the Instrumental Approach, we describe next how the tool can be instrumented in this situation, once assumed that responding students will build an instrument from the “calculation and graph” tool, within the item environment, in order to solve the task (Trouche, 2005, p 272). Instrumented action schemes are organized around the core elements that follow:

1. Knowing the difference between input and output in a contextual use of a function as a model.
2. Understanding that the tool returns unique heights (outputs) for both trees (numerical and graphical representations) when choosing a number of month (inputs).
3. Entering a number of months within the domain [0 ; 600].
4. Comparing outputs either in the numerical or graphical representation. Validating by linking to the real life situation.
5. Decide on the next number of month to enter considering the comparison to the previous one.
6. Iterating the process.

This type of instrumentation is linked to an operational approach of the concept of functions.

As mentioned earlier, the pencil can also be used in order to get a continuous model on the domain or part of it. Students can link points together using it. This step is an intermediate towards the primary task. The following core elements participate to instrumented action schemes using the pencil as well as the “calculate and graph” tool and are characteristic of a structural understanding of the concept of function.

1. Understanding the growth phenomenon is a continuous one. Hence, the functions modelling it are continuous.
2. Assuming both functions will be strictly increasing.
3. Using the pencil to link consecutive points together.
4. Decide on the next number of month to enter considering line intersection
5. Going back to numerical values to aim at accuracy

Of course one might operate composite instrumented action schemes, mixing the use of both principal tools. For example, one can use the pencil to sketch continuous graphs (potentially after trying 100 and or 200 months to get a more complete view of the graphs shapes), or rely on points’ colours and choose 400 months in the first tries and then use a trial and error strategy to aim precisely at the target with the “calculate and graph” tool.

Interactions at stake within the item are principally addressing the two different representations of the functions: numerical and graphical. When students use the "calculation and graph" tool, the feedback is given in both representations as new numbers in the table of values and two points on the graph. Students are consequently relieved from having to convert one representation into another. The colours of numbers and points related to the same function match so students can more easily interpret the feedback. This important interaction participate to students' reflection towards formulating the problem into both representations and then compare results (for example using colours' inversions in the graphical display) to either conclude or decide on other tries to make. Besides it can also participate to validate or invalidate tries that are outside the domain or very far from target.

LOGDATA ANALYSIS

CEDRE's interactive items, "tree growth" among them, have been experimented in a pilot test in May 2017 with a sample of 3 000 Grade 9 students per item. Students' digital traces have been recorded in log data files. As log data files contain a very large amount of data and in order to aim at interpretable results as well as to avoid noisy signals, variables of interest were defined, as a result of the *a priori* didactical analysis. They could potentially lead to build a model able to explain success or failure to the task considering either the operational or structural character of functions used by students. The main variables used in the analytical models are the following: Month list length, First input between 200 and 600, Number of alternating within the month list, Time spent on the item, Distance between first input and target value, Distance between the second input and the target, Distance between the last input and the target, Standard deviation of the month list, Target value is in the month list, Pencil use.

Classical machine learning analysis was then applied to the logdata. Unsupervised learning first aimed at clustering the sample. DBSCAN and k-means algorithms have principally been used at this stage. Supervised learning was also implemented to determine the predictive power of students' achievement of the model and estimate the weight of the variable in the prediction. Mostly logistic regression and random forest algorithms were used in the supervised learning stage. Mean values of the most important variables were then calculated for each of the clusters issued from the unsupervised learning.

RESULTS AND DISCUSSION

Clustering analysis distinguishes 4 clusters corresponding to 4 different students' profiles. Each cluster's size represents 25 % of the responding students. Non responding students (33 % of the sample) have been excluded from analysis as they did not use any of the available tools and did not input any number in the response box. It is to be noted that if the proportion of non-responding students is high, it is not higher than in other LSA such as PISA (OCDE, 2014). Figure 2 illustrates DBSCAN clustering results as well as important variables values for each cluster.

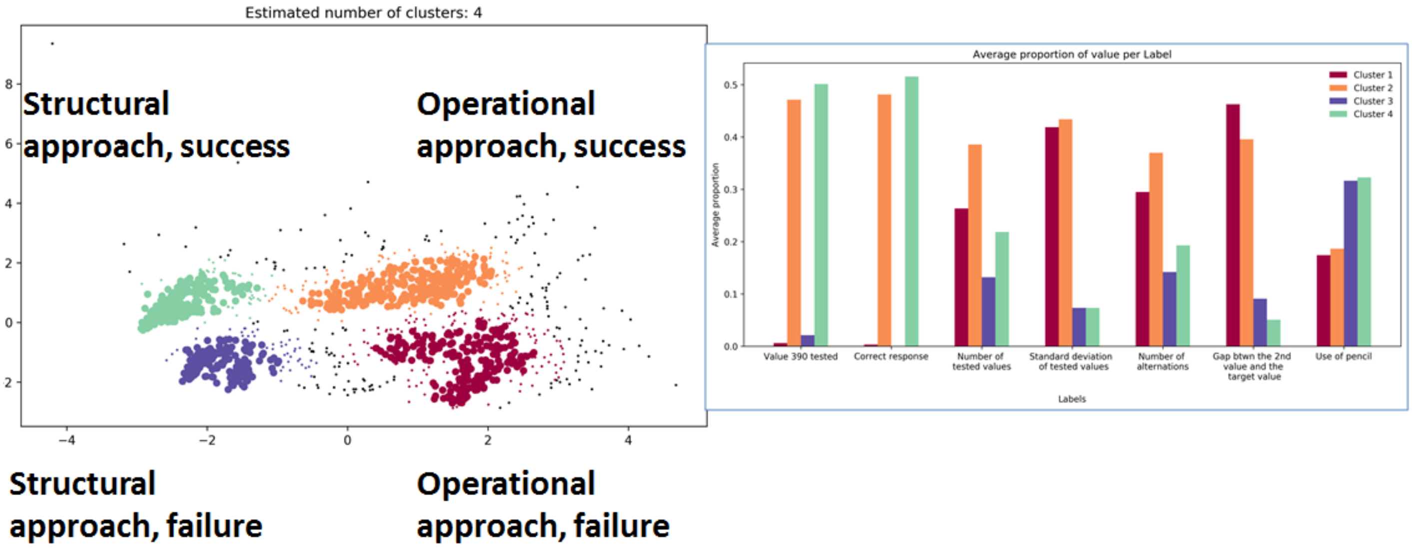

Figure 2: dbscan results and explanatory variables

Two profiles (green and orange in the graph) achieved the task. The other two (blue and red) correspond to students who failed. Apart from the obvious variable “value 390 tested”, no other variable allows to discriminate between success and failure. This result is disappointing in the sense that the model cannot explain students’ achievement to the item, which is one of our research questions. However, variables of interest in the *a priori* analysis allow describing profiles along another dimension than achievement. Two clusters (orange and red on the figure 2) show a large number of inputs, a large number of alternating inputs, a first input away from target and a large distribution of inputs. These characteristics allow us to interpret that students from these groups preferred a “trial and error” solving strategy, approaching the underlying concept of function in its operational aspect. Half of them achieved the task successfully, the other half did not. On the other hand, the other two groups share a different and opposite profile description according to the same variables: a small number of inputs, a small number of alternating inputs, a first input close from target, a narrow distribution of inputs. Moreover this second category of students used the pencil more often, altogether identifying solving strategies related to the structural aspect of functions. Alike the first two groups, the structural approach led to failing the task for half the students favouring it.

Hence, if the didactical and analytical models used in this study could not help us explain grade 9 students’ achievement to an interactive mathematical task in a computer based summative assessment, they could nevertheless help us identify two solving strategies well known in didactics literature. The potential usage of such result in diagnostic or formative assessment is obvious, this type of information being very valuable from teachers’ perspective in order to address students’ specific needs.

REFERENCES

Brousseau G. (1998), Théorie des Situations Didactiques, La pensée sauvage

Drijvers, P. and al. (2012). Tool use and the development of the function concept: from repeated calculations to functional thinking. International Journal of Science and Mathematics Education, 10, 1243-1267.

Kilpatrick, J. and al. (2001). Adding it up: Helping children learn mathematics, Washington, National Academy Press, p. 115-135.

Laborde, C. (2018). Intégration des technologies de mathématiques dans l'enseignement, in Dorier, J-L. et al, Enseigner les mathématiques. Didactique et enjeux de l'apprentissage, (p. 336-366) Belin

OECD (2013), PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy, pp. 43-44. OECD Publishing.

OCDE (2014), Résultats du PISA 2012 : savoirs et savoir-faire des élèves. Performance des élèves en mathématiques, en compréhension de l'écrit et en sciences, vol. 1, Paris, OCDE.

Rabardel, P. (1995), Les hommes et les technologies; approche cognitive des instruments contemporains. Armand Colin, pp.239, 1995

Robert, A. & Rogalski, J. (2005). A cross-analysis of the mathematics teacher's activity. An example in a French 10^{th} grade class. *Educational Studies in Mathematics, 59*, 269-298.

Roditi, E. & Salles, F. (2015) Nouvelles analyses de l'enquête PISA 2012 en mathématiques. Un autre regard sur les résultats In Education & Formations n° 86-87, pp.235-257, MENESR, DEPP

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. Educational Studies in Mathematics, 22, 1–36.

Stacey, K. & Wiliam D. (2013), "Technology and Assessment in Mathematics", in Clements M. A. (Ken), A. Bishop, C. Keitel, J. Kilpatrick and F. Leung (eds.), Third International Handbook of Mathematics Education, (pp. 721 – 752) Springer.

DOI: 10.17185/duepublico/70768
URN: urn:nbn:de:hbz:464-20191119-172649-8