Examining the findings of international studies from an educational perspective
The case of mathematics in PISA2012

Antoine Bodin

Instituts of Research on Mathematics Education - France

International Consultant, Curriculum and Assessment

Former Member of the Math Expert and Advisory Groups of TIMSS and PISA

Raphaël Couturier

Professor, Université de Franche-Comté, CNRS,

Institut FEMTO-ST, France

PISA Conference Dubrovnik, Croatia
29-30 November 2024

Communication plan

Rationale for the research (Why?)

Objectives of the research (What?)

Context and initial observations

Research methodology

Some results

Possible lessons for PISA participating countries

Rationale for the research (Why?)

PISA produces a considerable mass of data

widely used :

by managers and executives of educational systems,

by the expert teams in charge of developing curricula.

Sometimes useful for teacher trainers,

but not very useful directly for teachers and classroom practice.

In this research, we wanted to explore the possibilities to make better use of PISA data for mathematics teachers and teaching practices.

Objectives of the research (What?)

The project

Analysis of the results of the PISA 2012 survey on mathematical literacy

- Using the raw data
- From an international perspective
- From an educational and didactical point of view

Our research questions

- Beyond differences in achievement between countries, can PISA reveal significant differences in the students' cognitive organisation?
- Can Implicative Statistical Analysis (SIA) help to study this question?
- Can this type of research help teachers better understand students' cognitive behaviour?

Context and initial observations

The context is not very conducive to didactic research:

- With rare exceptions, PISA cognitive items are not disclosed.
- Not all pupils take the same test booklets, and therefore not the same items.
- Not all countries use all the proposed items.

For our research, we mainly used the raw PISA cognitive data (PISA2012lite) and the PISA 2012 Compendium for responses to cognitive items.

Since the 1980s, a method of analysis has been developed in mathematics didactics (and then extended elsewhere): Statistical Implicative Analysis (SIA). Its aim is not to classify students according to their level but simply to gain a better understanding of how they organise their knowledge and skills.

We therefore thought that this method might shed light on some of the questions posed by PISA.

About item difficulty

The difficulty of a test item is not an intrinsic notion (depends on the group of students tested).

However, PISA takes precautions to ensure fairness (i.e. Differential Item Functioning index).

So, the Item's difficulty order should be roughly the same across countries.

First observation:
We have compared the order of success for the 84 maths items shared by our eight countries by calculating the number of inversions needed to skip from the OECD order to this order for each country.

About item difficulty 2

….First observation :
Comparing the order of success for the 84 maths items shared by our eight countries by calculating the number of inversions needed to skip from the OECD order to this order for each country, we obtain :

Number of inversions in success order							
France	**Germany**	**Italy**	**Finland**	**Japon**	**Greece**	**Singapore**	**Korea**
247	**214**	**257**	**258**	**423**	**237**	**607**	**368**

Some of the inversions observed concern minor differences, but many show significant differences in scores.
For instance (range from less success to higher success):

PM949Q02T	JAPAN	FINLAND
Success	72%	13%
Rang	61	7

PM420Q01T	France	KOREA
Success	48%	28%
Rang	44	10

PM949Q02T concerns the interpretation of an incorrectly coded geometric figure
PM40NQ01T concerns the comparison of two series of whole numbers

Note that the order of scores for Asian countries is further from the OECD order than that of European countries.

Research methodology

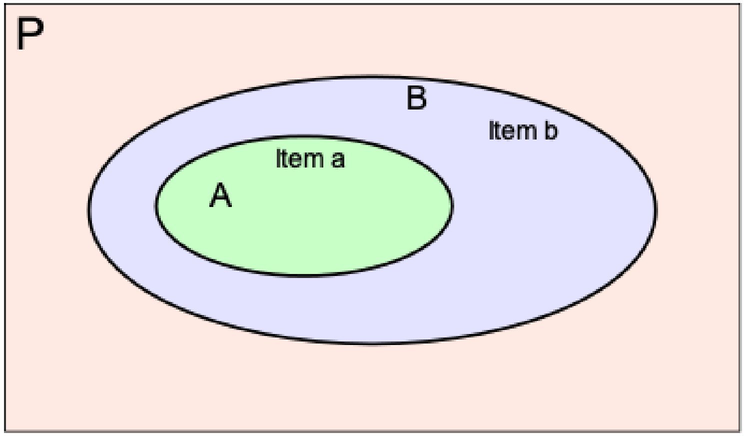
Statistical Implicative Analysis (SIA)

Two items, a and b, being administered to a population P.
Let a (resp. b) be the score obtained by a (resp. b), and let a ≤ b.
Let A (resp. B) be the group of students who have passed item a (resp. item b).

Two cases:

Case 1:
A is included in B ($A \subset B$)
It can be translated as « success in item a implies success in item b ».

a $\Longrightarrow$ b in the ordinary sense of logic.



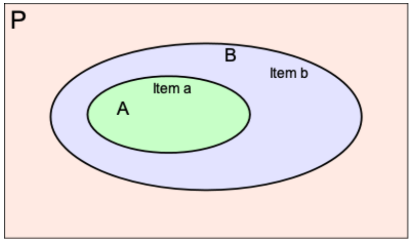
Success in a is a sufficient condition for success in b.
Success in b is a necessary conditin for success in a.

Research methodology 2

Statistical Implicative Analysis (SIA)

Two items, a and b, are administered to a population P.
Let a (resp. b) be the score obtained by a (resp. b), and let a ≤ b.
Let A (resp. B) be the group of students who have passed item a (resp. item b).

Two cases:
Case 1: A is included in B ($A \subset B$)
It can be translated as 'passing item a implies passing item b'. a $\Longrightarrow$ b in the ordinary sense of logic.



Case 2: A intersects B ($A \cap B \neq \emptyset$)

If 'a and not b' is small in front of A, we tend to think that a ALMOST implies b (fuzzy logic).

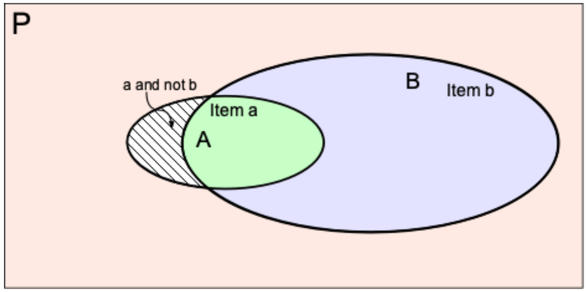
SIA modelises this intuition probabilistically by defining an intensity of implication linking a to b.

This intensity, noted φ (a , b) belongs to the interval [0 ; 1] and is not symmetrical.

For a given study, we choose a minimum intensity threshold (generally .95 or .99).

For a study involving large amounts of data, the intensity of implication is associated with a confidence index.

The PISA study was carried out for an implicative intensity threshold equal to .95 and a confidence index of .80.



Research methodology 3

Statistical Implicative Analysis (SIA)

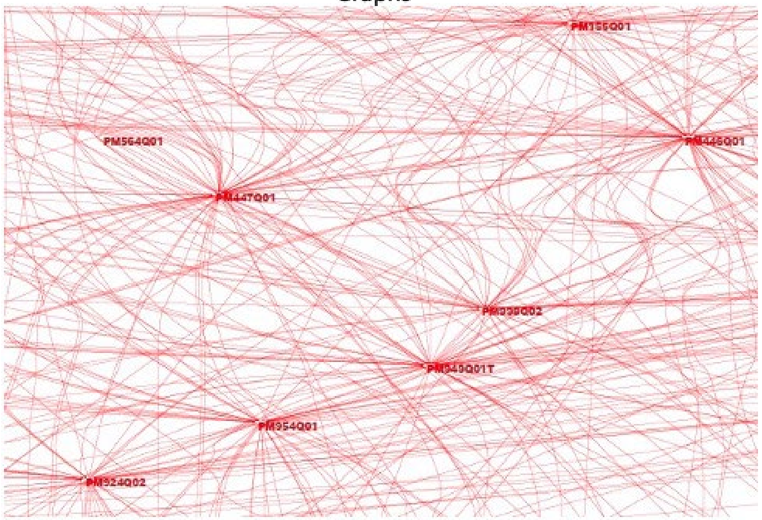
Given a square matrix of data crossing individuals and item results (binary data), SIA's Rchic software produces an implication matrix such that the intensity of implication φ (a , b) lies at the intersection of row a and column b.

	PM00FQ01	PM00GQ01	PM00KQ02	PM033Q01	PM034Q01T	....
PM00FQ01	0,00	0,88	0,83	0,95	0,90	....
PM00GQ01	1,00	0,00	0,97	0,83	0,96	....
PM00KQ02	0,96	0,90	0,00	0,77	0,93	....
PM033Q01	0,84	0,63	0,62	0,00	0,73	....
PM034Q01T	0,89	0,81	0,82	0,87	0,00	....
PM155Q01	0,79	0,68	0,56	0,81	0,75	....
PM155Q02D	0,89	0,72	0,57	0,82	0,80	....
PM155Q03D	0,99	0,88	0,85	0,99	0,94	....
PM155Q04T	0,86	0,63	0,78	0,62	0,93	....
PM192Q01T	1,00	0,79	0,76	0,96	0,97	....
PM273Q01T	0,88	0,87	0,77	0,85	0,95	....
PM305Q01	0,92	0,82	0,78	0,58	0,77	....
PM406Q01	0,98	0,93	0,96	0,83	0,87	....
....	....	....	....	....	....	....

Matrice d'implication

Rchic associates a graph with this matrix, making it possible to follow the organisation of the successes observed in a group of students.

Graphs



A piece of graph of PISA2012 for France Implicative Intensity .95, confidence 80% 27

Matrices and graphs can be used to compare the behaviour of different groups of students about the same test.

Rchic : R package developed by Raphaël Couturier

Research methodology 4

Comparison of implicative matrixes (SIA)

We defined an indicator of matrix density defined by the following formula where n is the number of variables.

$$d(n,\lambda) = \frac{The\ number\ of\ intensities\ of\ implications\ greater\ than\ \lambda}{\mathrm{n(n-1)}}$$

We obtain the following results *(…. missing data for Finland)*

Fresholds: implication .95 ; Confidence . 80

France: 40,3 % of the 6972 couples of variables show intensity of implication greater than .95.

Allemagne : 37,4 % ….

	FRANCE	GERMANY	ITALY	SINGAPORE	JAPAN	GREECE	KOREA
FRANCE	**0.403**	0.269	0.208	0.234	0.210	0.181	0.237
GERMANY		**0.374**	0.200	0.238	0.215	0.172	0.238
ITALY			**0.261**	0.181	0.169	0.159	0.184
SINGAPORE				**0.363**	0.211	0.143	0.251
JAPAN					**0.302**	0.137	0.231
GREECE						**0.261**	0.150
KOREA							**0.363**

Figure 9 : Density of implications (> .95); Country and Inter-country.

66% of France links are shared with Germany (0.269/0.403)

72% of Germany links are also France links (0.269/0.374)

80% of Italien links are shared with France (0,208/0,261)

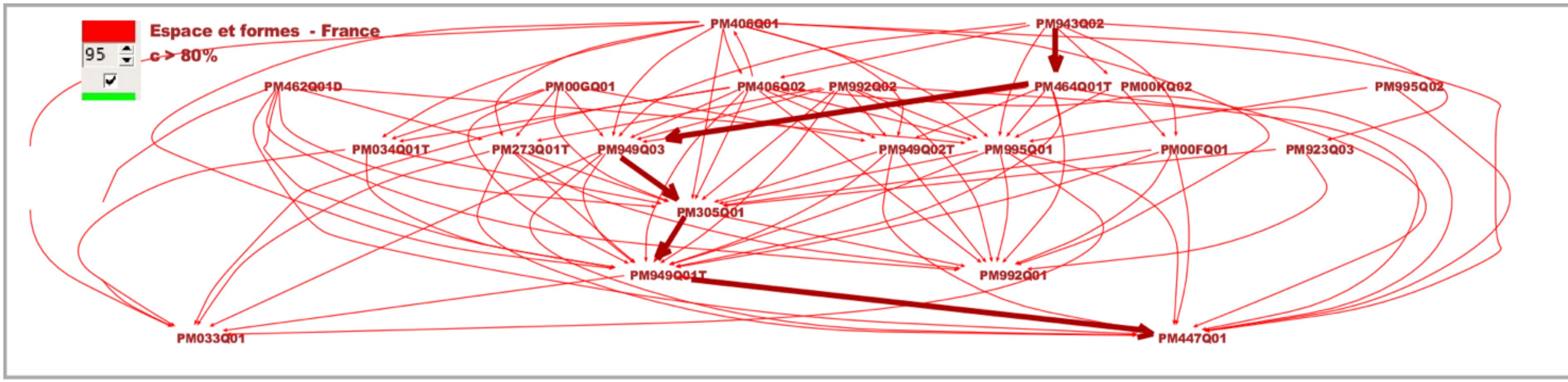
….

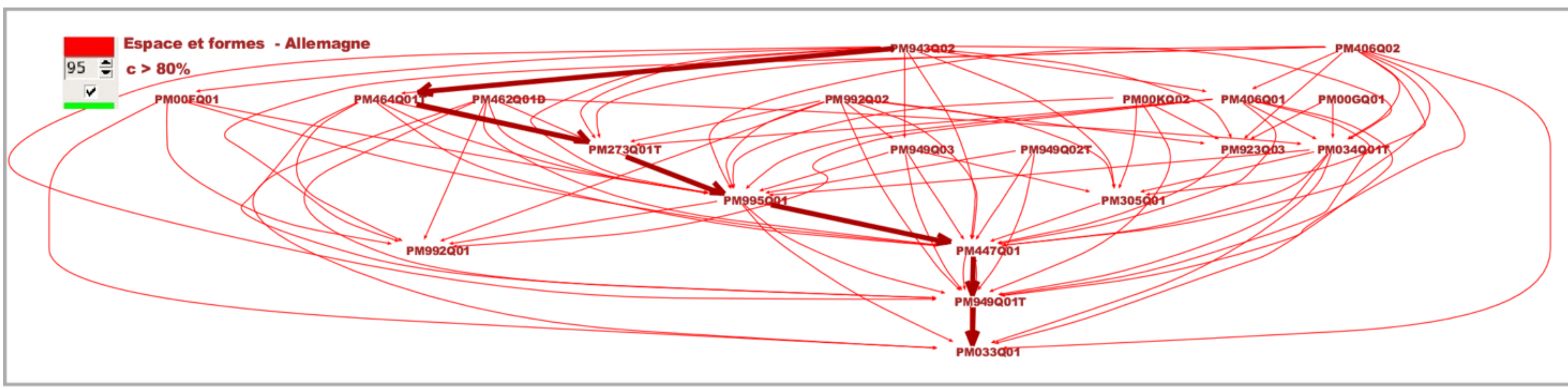
Smaller density indicates more knowledge dispersion

Research methodology 5

Comparison France- Germany

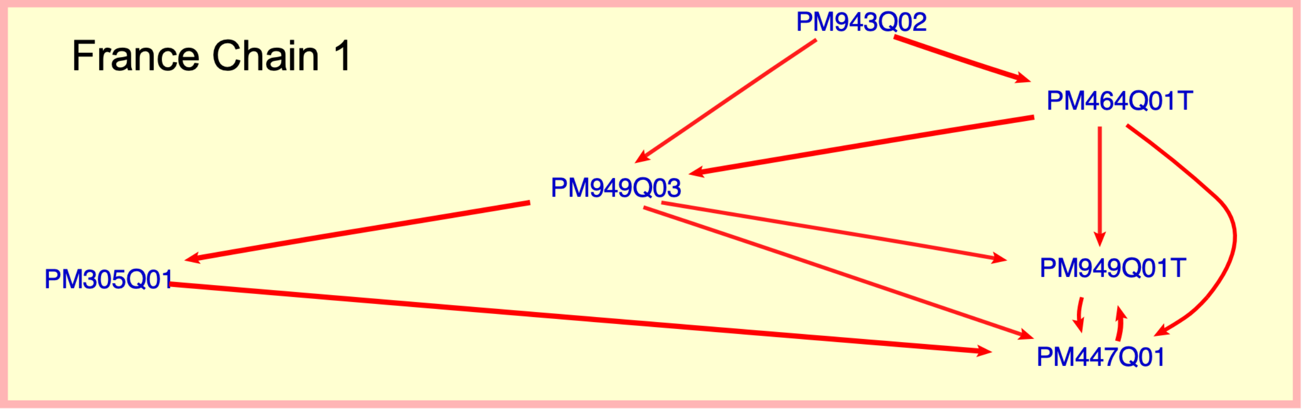
Graph of the 21 items Space and Shape for France and Germany with identification of paths of interest





Research methodology 6

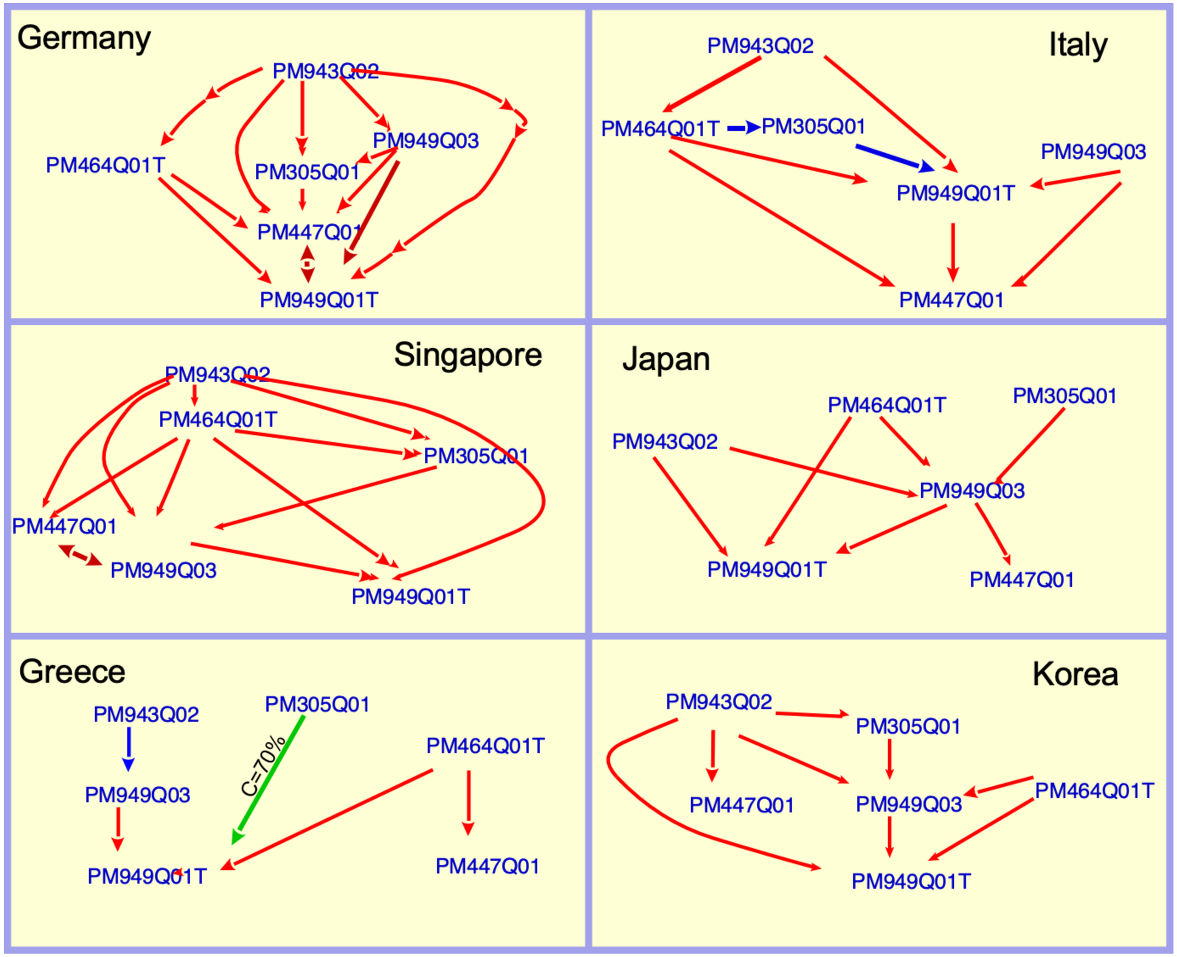
Comparison France- Germany



PISA Code	PM943 Q02	PM464 Q01T	PM949 Q03	PM305 Q01	PM447 Q01	PM949 Q01T
Name	Arches Q2	The Fence	Roof Truss Design Q3	Map	Tile Arrangement	Roof Truss Design Q1

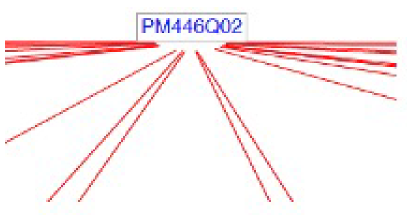
Research methodology 7

Comparisons same set of items in 7 countries



Key items

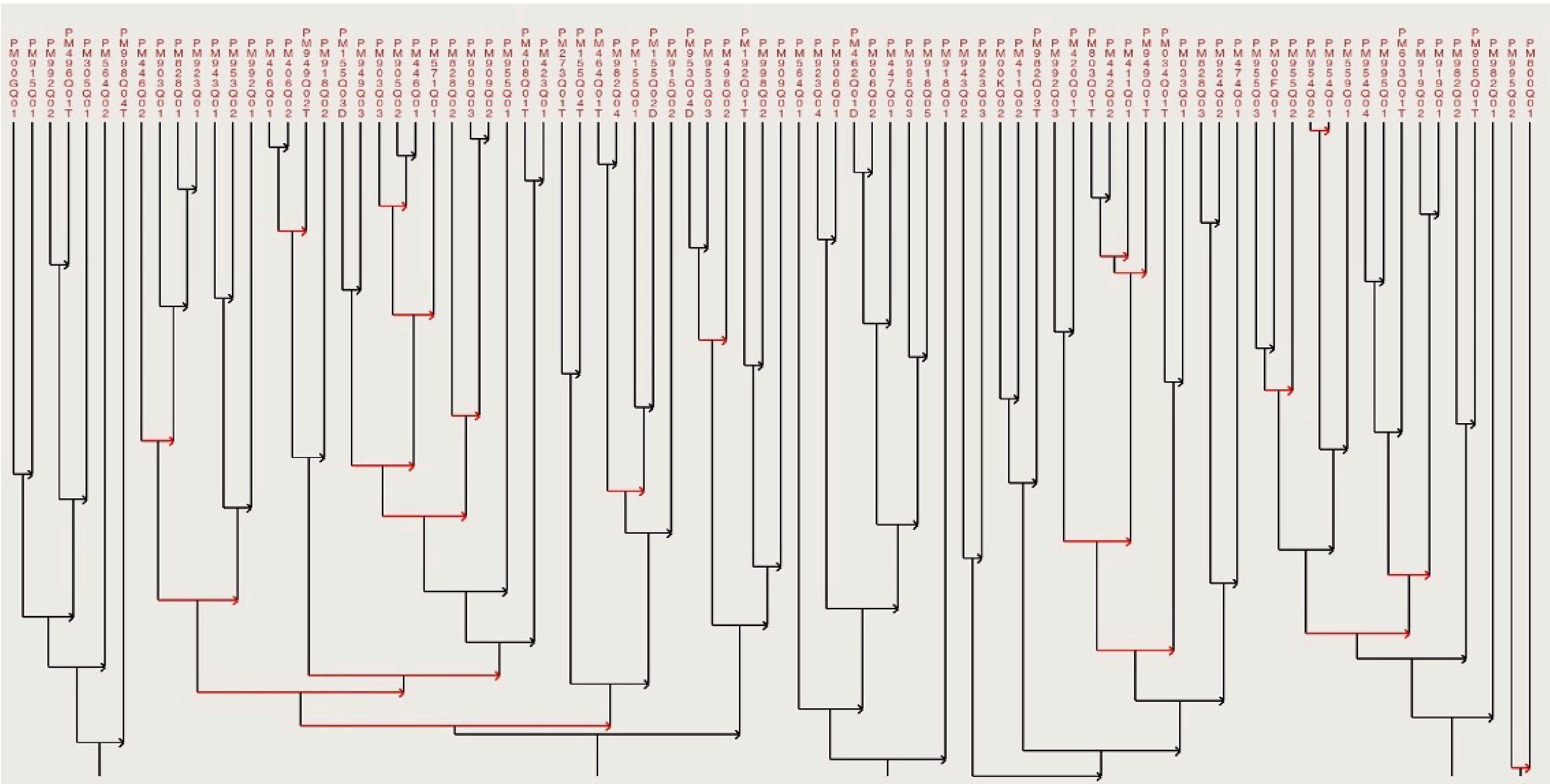
Implicative graphs produced by SIA are oriented.
Each arrow goes from a less successful item to a more successful item.
The graph shows : source items, pivotal items and wells items.

		
Source Item	Pivotal Item	Well Item

Some source, pivotal or well items are found common to several countries …

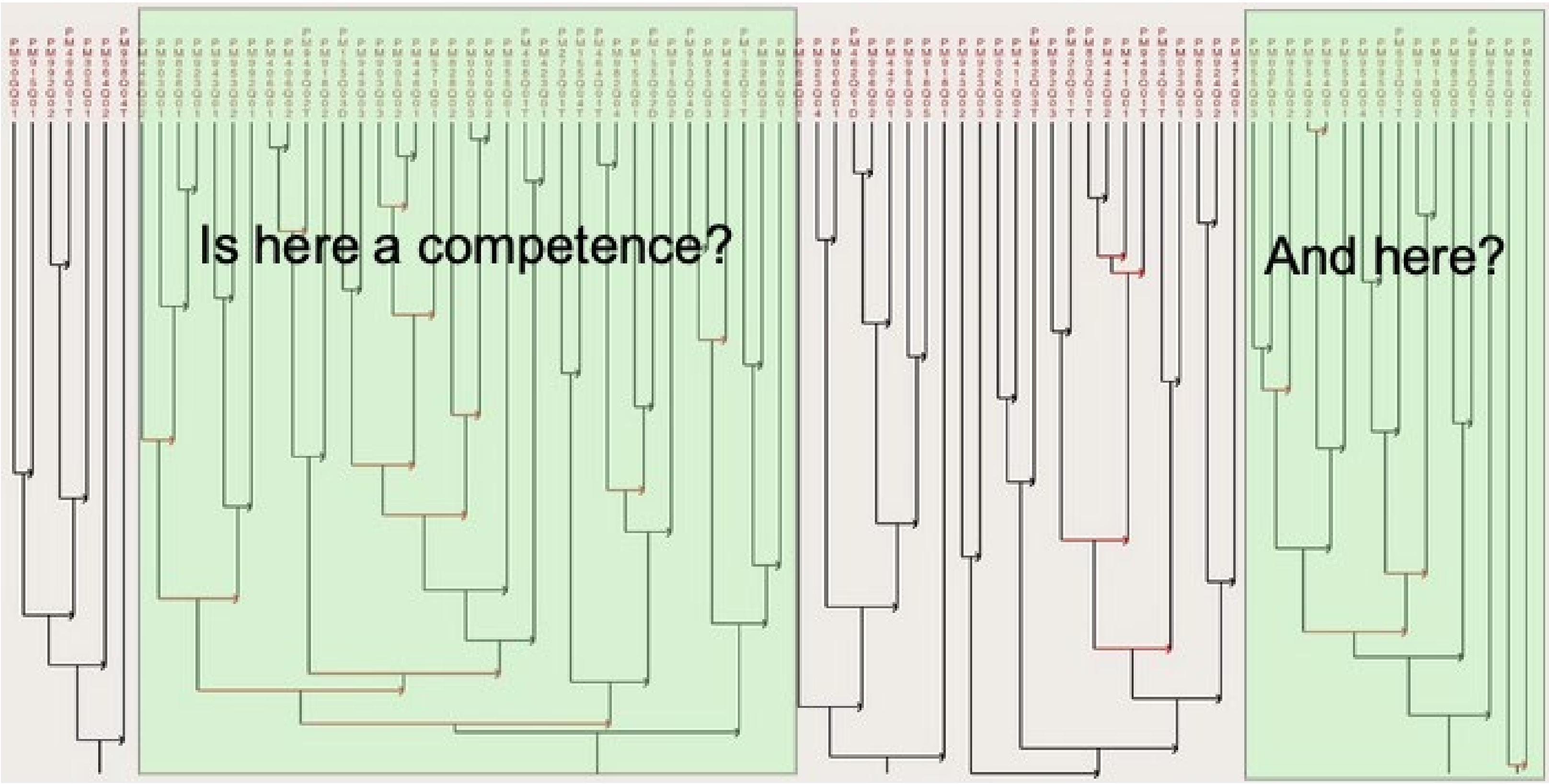
Cohesive Implicative Analysis

SIA also produces oriented implicative trees that allow for another analysis.
In particular, it helps to reconsider competency consistence.



This other way to present the implicative organisation of successes (here for France 84 items) opens ways for comparisons across countries.

Cohesive Implicative Analysis



If so for one country, are there the same across countries?

Some results

- The order of difficulty of the items varies across countries.
- Depending on the country, the results for the items are more or less related, reflecting a greater or lesser degree of cohesion (or dispersion) in students' results.
- Some items may be considered key items: their success may be considered necessary (or sufficient) for success in a significant number of other items.
- Items strongly related in one country are not necessarily so in another.
- We translate these observations by saying that the cognitive organisation of pupils varies from country to country.

Possible lessons for PISA participating countries

- Over and above differences in the overall level of mastery in mathematical literacy, the use of the SIA reveals differences in the cognitive organisation of pupils in different countries.
- The causes of these differences can be traced partly to cultural differences but also to curricular differences and differences in teaching practices.
- The observations made possible by SIA are of interest to teachers and teachers trainers. They provide a better understanding of the difficulties encountered by students and offer avenues for pedagogical action to overcome them.

Note that the ASI and the method presented here can be used for other PISA data (Science Literacy, Reading Literacy, etc.) as well as for data from other international studies.

References

Bodin, A. & Grapin, N. (2018), Un regard didactique sur les évaluations du PISA et de la TIMSS: mieux les comprendre pour mieux les exploiter. Mesure et évaluation en éducation, 2018, vol. 41, no 1, 67-96.

Bodin, A., Couturier, R. Khaled, H. (2019), Les mathématiques de PISA 2012, huit pays à la loupe de l'Analyse Statistique Implicative, actes colloque international ASI, Université de Franche Comté.

Bodin, A. (2008), Lecture et utilisation de PISA pour les enseignants. Petit x n°78, IREM grenoble, 53-78.

Suurtamm, C. & al. (2016), Assessment in Mathematics Education, ICME-13 Topical Surveys, DOI 10.1007/978-3-319-32394-7_1

Gras, R., Bodin, A., Couturier, R., Gregori, P. (2023), The Theory of Statistical Implicative Analysis. Or the Implausibility of Falsehood ... When the Exception Confirms the Rule. CRC Press.

Gras, R. Einoshin Suzuki, Guillet, Spagnolo, F. (Eds.) (2008) Statistical Implicative Analysis, SPRINGER.

Bodin, A. & Trouche, L. (2019), Penser PISA dans la perspective du développement de l'enseignement des mathématiques en Algérie.

antoinebodin@mac.com https://antoine-bodin.com/

raphael.couturier@univ-fcomte.fr. https://members.femto-st.fr/raphael-couturier/fr