HOW DO GRADUATEMATHEMATICS STUDENTS EVALUATE ASSERTIONS WITH A FALSE PREMISE ?

Janine Rogalski, MarcRogalski

Laboratoire Cognition & Activités Finalisées Université Paris8-CNRS

Institut Mathématique de Jussieu CNRS-Universités Paris 6 et Paris 7

Abstract

We present an empirical study focusing ondiagnosis of graduate mathematical students' logical ability. Patterns ofresponses on evaluating non-computable implications are identified and related to answers in hypothesis testing (Wason's tasks)."Stable" categories either oriented by logic or by pertinence are stronglylinked to success in the final exam. The category of "false logical"answers is the worse. Some issues about didactical consequences are discussed.

Introduction

Mathematical reasoning often requires touse assertions with uncertain or false premises. Such a situation isencountered, for instance, in proving the step in a recurrence, or in proofs needing "case by case reasoning", when the premise is not true for some setof parameters values. In order to master such situations, students must:clearly understand the concept of implication—involving or notinvolving quantifiers—, master the notion of counterexample, and be fully aware of the distinction between implication and equivalence. More, taking the negation or thecontraposition of a given mathematical statement requires to differenciatethe implicative link from its components: premise and consequent.

To our knowledge, there were fewempirical studies reported on this subject with regards to the litteratureon mathematical proof or on natural implication in experimental psychologyof reasoning. However, some epistemological or psychological analyses can be found on this point,which all insist on the conflict between everyday reasoning on falsepremises and mathematical logical reasoning, or which present thesesituations as paradoxical ones (for instance: Legrand, 1990; Johnson-Laird & Byrne, 1991;Deloustal-Jorrand, 1999; Politzer & Carles, to appear). In a grand tour onwhy students cannot master mathematical reasoning, Dreyfus (1997) stressedthe "shortage of research data" on students' explanations when reasoning in undergraduate mathematics. Infact, studies on reasoning emphasise global proof strategies, organisationand understanding (for instance: Balacheff, 1987; Coe & Rutven, 1994;Duval, 199; Fischbein, 1982). Amongresearches which focus on details on mathematical logic used indemonstration, Radford (1985) analyses implication by secondary levelstudents; Durand-Guerrier(1996) stresses the role of contingent assessments and the importance ofpredicats in logic, and not only propositions, for college students.

In mathematical curricula, logic isgenerally not taught as such, even if textbooks may present it as a tool oras an object at the university level (Deloustal-Jorrand, 1999). Theusefulness of logic in compulsory mathematics education is questionned for futurenon-mathematicians. At high-school or university levels, emphasis is put onmathematical contents (concepts and problem solving): the main idea is thatlearning to solve problems and proving statements is sufficient forlearning to master mathematical reasoning.Our claim is that it is probably unsufficient, because most of the proofsencountered in mathematical cursus, including at the university level, aremainly "going from stated as true to stated as true".

The present study focuses on what happenwhen such a process is disrupted.

The experimental study

Students and tasks

Students were graduate students (threeor four years after *baccalaurŽat*), candidates from theNorth Region of France for being mathematical teachers at the secondarylevel (107 students—59 girls and 48 boys). They answered a test about their global reasoningability at the beginning of their preparation to the national competitiveentry examination. The test duration (4 hours) was sufficient to answer allthe 20 questions. It was compulsory, answers were non anonymous, and students were said thatthe test aimed at informing the teacher on their collective logicalcompetence.

We will here analyse 7 non mathematical or elementary mathematical items,where the task was to assess the value of an implication. In 5 cases thepremise of the implication was clearly false; the two other items were theclassical Wason's selection task (Wason, 1966) and the Radford's version of it (Radford, 1988). The text ofthese items is given in annex, with their respective positions within theset of 20 questions. Through the forword, we intended to avoid contract effects, to obtain comments, and to avoidnon responses, as far as possible.

Data analysis

We will present the first step of dataanalysis, based on categories of responses, without taking into accountdetailed procedures.

Categories ofresponses

• Three of the questions concernedmathematical or material implications. The implication truthfulness was**non computable**. These items were coded as "*non-computable implication*". Premises were meaningful and even provoking. There were 4 types ofresponses: true; false; non sense (or variants); no answer.

• Another mathematicalimplication to be assessed was a **computable one**("*Hn⇒Hn+1*"**)**. In a previous part of the question, it was asked to prove the implication*via* a simple computation. It is coded "*computable implication*". The premise was without direct meaning forthe students.There were the same types of responses as for non-computable implications.

• Two items were two versions of**Wason**'s selection task, coded as "*Wason*" (W1 and W2). There were 4 types of responses: all correct; W1 correct &W2 non correct; W2 noncorrect & W1 correct; other or no answer. Two errors were particularlyinteresting from the point of view of the research of counterexamples:missing the non-Q card (with 7) in W1 and evaluating the procedure 2 as nonconclusive in W2. They were coded“non-Q missing”.

• The last item was a situationof **social contract** ("*The teacher'ssweets*"): "If P then I will do Q", then P is not satisfied and Q isnevertheless performed.Four types of responses were identified: logical true; modifying thecontext to make P true; contract not respected; other or no answer.

Pattern of responses inevaluating assertions with false premise

The second step of analysis consists in defining patterns of responses on the three similar items(non-computable implication), and expressing their relationship with theother ones, and with the success to the final mathematical competitiveexam.

Seven patterns were identified:

• **Logic** perspective, **stable** (LS): 3 "*true*"logical correct answers; 10 students.
• **Logic** perspective,**unstable** (LI): two "*true*" answers; 9students.
• **Pertinence** perspective, **stable**(PS): three "*non-sense*" answers denying pertinence to thequestion ; 9 students.
• **Pertinence** perspective, **unstable** (PI): two "*non-sense*" answers; 14students.
• **Non-conditional** answers (NC): three "*false*" answers (with or without explanation: "*because Pis false*"); 22 students.
• **Non-conditional** answers, unstable (NCI): two"*false*" answers; 23 students.
• Answers with **no-dominant** type of answer (SD); 20students.

Given the global results, theprobability to find a dominant pattern (all but SD) if students answeredindependently to each item would be quite lower than the observedpercentage, while the probability to findSD patterns would be higher.

Results

Global results

They are presented, for implications withfalse premise, in Table1, and for Wason's tests, in Table 2.

Table 1. Number of students given each typeof answer depending on the questions

	True	non pertinent	False	other
triangle	15	38	46	8
lights	13	33	51	10
labyrinth	29	24	48	6
Hn⇒Hn+1	57	17	20	13
*the teacher's sweets**	31	37	11	28

* For this item, the “non pertinentquestion” answer was dominantly "modifying the context in

order to make the premisetrue"; the answer “contract not respected” is put under“assertion false”.

The two first items were quite similar. Forthe “*labyrinth*” item, there were more“correct” answers and less “nonpertinent” ones: it could be possible that some of the students missed the premisefalsity because they “forgot” the rule of movement.

Table 2. Number of students given each typeof answer on Wason’s tasks

	Correct	non-Q missing	other errors	no answer
Wason 1	51	40	13	3
Wason 2	77	20	5	5

There were more correct answers with the Radford’s version (W2) thanwith the Wason’s selection task (W1). In fact, 43 students (40,2%)gave both correct answers; 34 correctly evaluate the procedures for testing the rule and made errors in the selectiontask, while only 8 were correct for W1 and made errors for W2.

Relationships betweentype of questions

Figure 1. Percentage of students with acorrect answer to the computable assertion, depending on their patterns ofanswers to non-computable implications.

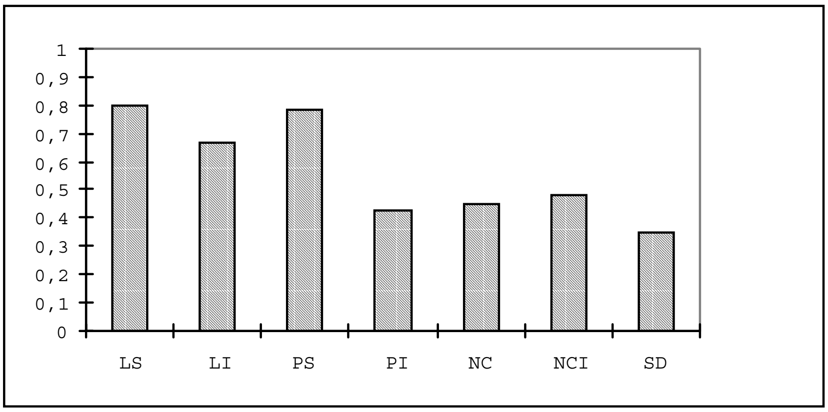
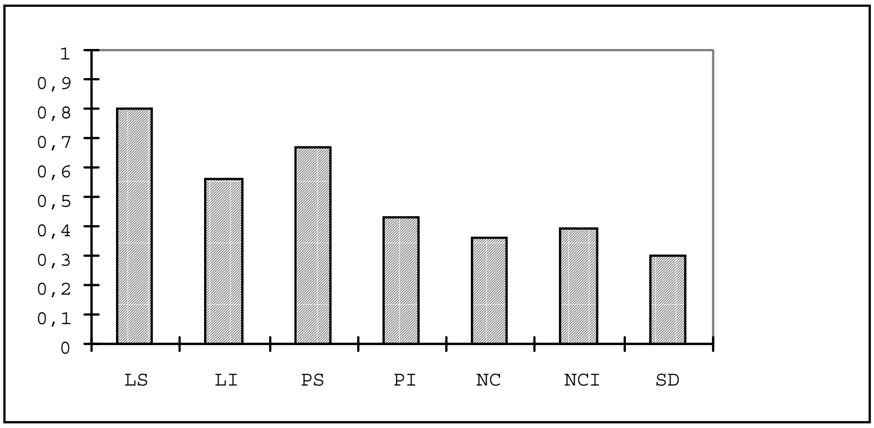


Figure 2. Percentage of students with correct responses to Wason's selection tasks,depending on their patterns of answers to non-computable implications.



Patterns of answers presented similarrelationships with the correct answer to the computable implication and tothe correct answers to the twoWason's tests. Both stable logic and pertinence perspectives (LS and PS)ensured success, while answers without dominant perspective (SD) led to theworst results.

However, a further analysis showedthat 80% of LS students succeeded both the Wason's tests and the computableimplication, while the other categories were all under 45% of success (17 %for NC, NCI and SD groups).

Mathematical and materialimplications versus social contract statements

Figure 3 presents the relationship betweenthe responses to non-computable implications and the logical answer to thesocial contract item.

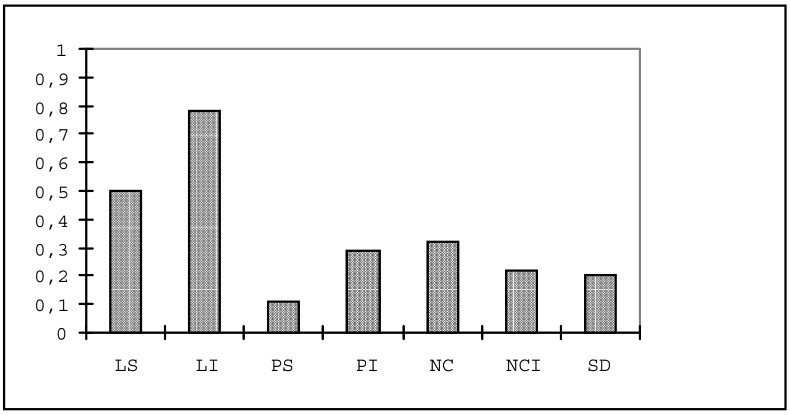


Figure 3. Percentage of students with a correct logical answer to the socialcontract question, depending on their patterns of answers to non-computableimplicatio ns.

Logical answers to the social contractquestion were low (less than 30%), except forstudents giving logical answers (stable or incomplete) to the mathematicaland material implications (globally: 63% of logical responses to the socialcontract item). The lowest percentage (10%) was observed for studentsanswering with a (stable) pertinence perspective (PS).

Reasoning about false premisesand mathematical success

The following table indicates the percentage(and number) of students passing the written test (first step of the competition), and the percentage of students whosucceeded in this competition (and entered a teacher training institute:IUFM).

Table 3. Results to the competitiveexamination

	National level	Regional level	_Tested students
Attending thecompetition	7780	638	107
Passing the written test	28,4% (2212)	28,8% (184)	38,3% (41)
Fully succeeding	12,4% (968)	12,2% (78)	18,7% (20)

With regards to their success in the national competitive examination, thestudents in our study were above the national and regional levels. Thismust be taken into account for evaluating the generality of theresults.

Figure 4. Percentage of students successing the written final exam, depending ontheir patterns of answers to non-computable implications.

Figure 4 presents the percentage of students passing thewritten test depending on their patterns of answers to non-computableimplications.The best results were observed for students responding from a stableperspective of logic (LS) or pertinence (PS) (70% success); the worstresults concerned the students with coherent non-conditional responses (NC:less than 10% success). Students presenting variability in their answers (LI, PI, NCI and SD) wereslightly above the national mean (between 30% and 50%).

Discussion and conclusion

Students' answers concerning assertionswith false premise and hypothesis testing, and students' success to thefinal mathematical show interesting relations. Both students centered onlogicand on pertinence presented similar success. Only students who focus onlogic answer logically to the social contract question. The model:"implication is false when premise is false" appears to be the worse forhypothesis testing and mathematical success.

The results on Wason's tasks have to be underlined: our results not onlyconfort Tweney and Yachanin results (1985) on the fact that scientistsrationally assess conditional inferences better that non scientists, butalso that graduate mathematics students can do better than the university physicians members tested byWason.

These diagnosis results are partiel elements of a more complete study (onthe 20 questions of the test) which might enable us to analyse the variousdifficulties —and successes— of mathematical graduatestudentsin logical processes used in mathematics.

The long term purpose of understandinghow students process at a fine grained level of their mathematical logic isto identify what could be efficiently taught in this field, while keepingin mind the shared goalexpressed by Thurston: the most important in not to prove but to understandmathematics.

The main issue is to improve the logicaltool, in agreement with Hanna and Jahnke (1993) who defend the positionthat "a curriculum which aims to reflect the real role of rigourous proof in mathematics must present it as an indispensable toolin mathematics rather than the very core of that science" (p. 879).

References

Balacheff, N. (1987). Processusde preuve et situations de validation.*Educational Studies inMathematics, 18*, 147-176.

Coe, R., & Ruthven, K. (1994). Proof practices and constructs of advancedmathematical students. *British Educational Research Journal, 20*,41-53.

Deloustal-Jorrand, V. (1999). L'implication. Quelques aspects dans lesmanuels et points de vue d'Žl□ves-professeurs. *Petit x, 55*,35-70.

Dreyfus, T. (1997). Why Johnny can't prove. *Educational Studies inMathematics*, *28*, 85-109.

Durand-Guerrier, V. (1996). *Logique et raisonnement mathŽmatique*.Th□se de Doctorat. UniversitŽ Claude Bernard Lyon 1.

Duval, R. (1991). Structure du raisonnement dŽductif et apprentissage dela dŽmonstration. *Educational Studies in Mathematics*, *22*,233-261.

Fishbein, E. (1982). Intuition and proof. *For the Learning ofMathematics, 3*, 9-24.

Hanna, G., & Jahnke, H. E. (1993). Proof and application. *EducationalStudies in Mathematics*, *24*, 421-438.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hove andLondon, UK: Lawrence Erlabaum Associates.

Legrand, M. (1990). "Circuit" ou les r□gles du dŽbat mathŽmatique.In *Enseigner autrement les mathŽmatiques* (pp. 129-161). Paris :Commission Inter-IREM UniversitŽ.

Politzer, G., & Carles, L. (to appear). Belief revision and uncertainreasoning. *Thinking and Reasoning*.

Radford, L. (1985). *InterprŽtation d'ŽnoncŽs implicatifs ettraitements logiques. Contribution ˆ la faisabilitŽ d'un enseignementde la logique au lycŽe*. Th□se de l'UniversitŽ de Strasbourg.

Tweney, R.D., & Yachanin, S.A. (1985). Can scientists rationaly assessconditional inferences? *Social Studies of Sciences, 15*, 155-173.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.). *New horizons inpsychology.* (Vol.1). Harmondsworth: Penguin.

Annex. The logical test proposed to students consisted on 20 questions (11mathematical ones, nine requiring no mathematical knowledge). Among them,seven asked to evaluate the value of an assertion or of a rule. The text ofthese questions is given below, with their number. The test began with the following forword: "Mathematicalsentences, assertions, in the following exercices are often expressed in anaive, not formalised form, asit is done in an everyday mathematical text, or even in everydayconversation. It is deliberate, and you have to feel free in your answers,which may (must !) include many comments, and even propose "this questionis a stupid one!". In italics the contextof the analysed question.

Computable assertion with false premise

I. *Given* $\lambda \in$ *″,let be* $(u_n)_n$ *the sequence defined by :*$u_0 = \lambda$*,* $u_{n+1} = 2u_n + 1$*.*

Let H_n *be the assertion "* $u_n{}^2$ *\f(*2^n*;3) - 1".*

(a) Is the implication " $H_n \Rightarrow H_{n+1}$" true for some n ? for every n ?

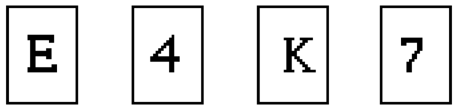
(b) Compute explicitly u_n as a function of λand n [one may write $u_n = v_n - 1$].

(c) Show that if λ > - \f(2;3) alltssertions H_n are false.

(d) If λ = 10, what can be said about the assertion" ∀n $H_n \Rightarrow H_{n+1}$ " ?

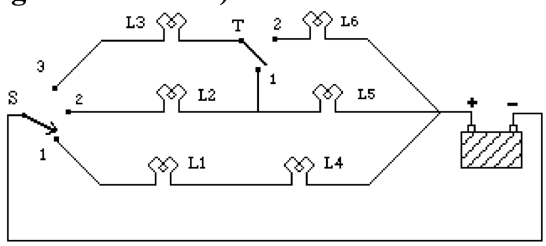
Non computable assertion with false premise: item 1
III. What do you think about the truthfulness of the followingassertion:
"Every not flat triangle of the plane, whose mediatrices are notconcurring, is an equilateral triangle" ?

Wason's selection task (W1)
IX. Cards aregiven, with a letter on one face and a number on the other one. One musttest the possible rule: "behind a vowel there is an even number". For thispurpose, a sample of 4 cards is disposed on the table; what is seen isrepresented just on the right. What card(s) is/are to be turned on in order to know is the rule is confirmed onthis sample?".



Radford's version of Wason's task (W2)
XV. A set of numbered balls are in an urn. Balls are blackor white. The following question is "do all white balls in the urn get aneven number?"
Four procedures are considered for answering the question:
procedure 1 : balls with an even number are taken out of the urn; thentheir colour is considered;
procedure 2: balls with an odd number are taken out of the urn; then theircolour is considered;
procedure 3: white balls are taken out of the urn; then their numbers areconsidered;
procedure 4: black balls are taken out of the urn; then their numbers areconsidered.
For each of these 4 procedures, chose among the two options:
(a) the procedure will certainly enable me to answer the question;
(b) there is a chance that the procedure does not enable me to conclude.

Non computable assertion with false premise (Legrand's Circuit): item2
XVII. An electric circuit consists on six identical lampsdenoted L1, É L6, and two switches S and T; S may take three positions: S1, S2, or S3,and T may take two positions: T1 or T2.



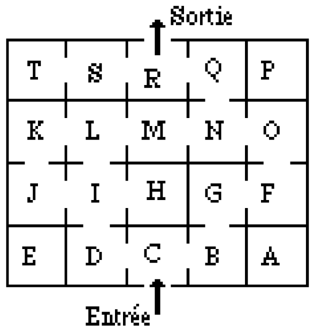
(a) What may be said about the truthfulness of thefollowing assertion: "if L1 is turned on or if L6 is on, then L3 is on orL4 is on"?
(b) What may be said about the truthfulness of the followingassertion: "If L1 is on and if L3 is on, then L2 is on and L5 is noton"?

Social contract question
XVIII. In a primary school, the teacher gives a problem to the pupils and says:"search this problem at home; tomorow, if somebody was able to solve it, Iwill give you sweets". The following day, no pupil has been able to solvethe problem. The teacher distributes sweets to the pupils. They protest: "It is not just, we were not ableto solve the problem, we get no right to sweets!". Mocking, the

teacher answers that she perfectly respected the contract... What are your comments about this story?

Non computable assertion with false premise (Durand-Guerrier'sLabyrinth): item 3

XIX. Somebody, called X came through the labyrinth (on theright), from the entrance (entrŽe) to the issue (sortie), without going twice through the same door. Roomsare called A, B, É, T.



For each of the seven following sentences, say if it is true,it is false, or if it is impossible to know:

(1) X went through T ;
(2) X went through N ;
(3) X went through M ;
(4) if X went throughO, then X went through F ;
(5) if X went through K, then X went through L ;
(6) if X went through L, then X went through K ;
(7) if X went through S, then X went through T.