Traitement de la validité de l'implication par des étudiants, corrélations avec leurs performances mathématiques, liens avec diverses questions de psychologique cognitive

Janine ROGALSKI, Marc ROGALSKI

Laboratoire Cognition et Usages, Université Paris 8
Laboratoire Paul Painlevé, Université de Lille 1-CNRS

Introduction

Nous présentons un ensemble de résultats d'une étude faite sur deux groupes d'étudiants licenciés de mathématiques entrant dans la préparation au CAPES de mathématiques à l'Université de Lille, la plus nombreuse en septembre 1999 (test 1: 107 étudiants) et la deuxième en septembre 2001 (test 2 : 71 étudiants).

La motivation de l'étude réside dans l'importance, pour des enseignants de mathématiques en collèges ou lycées, de savoir distinguer dans les “erreurs de raisonnement” des élèves celles qui proviennent directement de la mauvaise compréhension des mathématiques enseignées, et celles qui sont liées à un maniement erroné de la logique en œuvre en mathématiques ; par exemple, un certain nombre d'élèves et d'étudiants pensent spontanément, en actes, que P⇒Q et nonP⇒nonQ, "c'est pareil", ou que la négation — souvent d'ailleurs dénommée “contraire” — de "∀x P(x) est "∀x non-P(x)", ou distinguent mal une implication et sa réciproque : “ dire P⇒Q et dire Q⇒P c'est dire la même chose” (une étude extensive sur de bons élèves du niveau de 4ème — 8th grade — met en particulier ce point en évidence :Küchemann & Hoyles, 2002). L'origine de ces erreurs est souvent plus cachée, et nous faisons l'hypothèse que les enseignants ne seront vraiment aptes à les détecter et à proposer des remédiations éventuelles que si eux-mêmes ont des idées claires sur le sujet (cf. Durand-Guerrier, 1996 ; Durand-Guerrier et al., 2000). Ce sont donc ces idées que nous nous proposons d'étudier sur ces futurs enseignants.

L'étude reprend d'abord en partie des résultats publiés par ailleurs (Rogalski & Rogalski, 2003), et nous renvoyons le lecteur à ce texte pour plus de détails. Dans un premier temps, nous nous sommes concentrés sur l'implication, et en particulier sur les divers modes de traitement de la validité de l'implication que ces étudiants "avancés" pratiquent. Nous avons laissé de côté pour l'instant l'étude de "modes de raisonnement mathématiques" plus globaux, que nous reprendrons éventuellement plus tard : nous travaillons plutôt sur "le pas de raisonnement" au sens de R. Duval (1991), mais en prenant en compte aussi bien des pas de raisonnement sans quantificateurs, qu'avec un – ou plus d'un – quantificateur (voir à ce sujet les travaux de V. Durand-Guerrier et ceux de M. Legrand).

On présente ensuite les relations entre modes de traitement de la validité de l'implication et traitement de questions mathématiques, posées dans le cadre des mêmes tests. Les résultats globaux du concours du CAPES sont également considérés de ce point de vue, pour évaluer la pertinence d'une étude centrée sur l'implication. De multiples travaux ont montré le rôle de la preuve dans l'enseignement des mathématiques (voir la revue de questions de Hanna, 2000), ainsi que les relations entre compétence logique et compétences mathématiques, qu'il s'agisse du langage mathématique (différant de l'usage de la langue naturelle), de la disponibilité limitée des contre-exemples (Benbachir & Zaki, 2001).

Nous pensons en effet que le "pas implicatif" en mathématiques est une forme d'inférence constamment présente dans tous les raisonnements heuristiques et toutes les argumentations en jeu dans une recherche mathématique, bien avant la mise en forme rigoureuse de la démonstration. Anticipant sur les résultats qui suivent, nous ajoutons qu'il faudra aussi s'engager dans une prise en compte du contexte d'ensemble de ces activités de raisonnement (niveau des sujets —élèves, étudiants, enseignants—, visée du travail mathématique, contenu mathématique et rapports des sujets à ce contenu mathématique, types d'implications en jeu, visée didactique elle-même).

Du point de vue de l'implication, nous nous sommes centrés sur les modes de traitement par les étudiants des implications à hypothèse fausse, ou "parfois fausses" lorsque les implications concernent des assertions quantifiées (les "hors-sujet", Legrand, 1990, p. 141, cf. *infra*). Ce faisant, nous nous démarquons fortement des très nombreuses études de psychologie sur le raisonnement et particulièrement la déduction, qui partent de propositions tenues pour vraies. De plus, nous nous sommes intéressés aux relations entre un ensemble de réponses à des tâches mettant en jeu l'implication, pour identifier des différences individuelles – des "profils" dans les modes de traitement de l'implication, et les cohérences ou contradictions entre réponses. Pour situer nos résultats par rapport à ceux apportés par la psychologie, nous avons introduit un "classique" des études sur la déduction : la tâche de sélection de cartes de Wason (1966), et la version qu'en a construite Radford (1985).

La dernière partie de ce texte fait un point sur le développement des approches de la psychologie cognitive sur la déduction, le raisonnement, et la rationalité. Ce tour d'horizon est évidemment schématique, vue la profusion des travaux sur ces thèmes (ainsi, une revue leur est maintenant consacrée: "*Thinking and Reasoning*").

I. Hypothèses, typologie des implications, "profils" de réponses

Détaillons maintenant les hypothèses qui sous-tendent l'exploitation empirique des données recueillies.

La première hypothèse que nous avons faite est qu'on peut avoir accès aux schèmes d'utilisation de l'implication présents dans la population étudiée et aux effets de ces schèmes en dépouillant les réponses des étudiants à des items qui, soit sont des questions de validation d'implications dont l'hypothèse est toujours fausse, soit attirent l'attention des sujets sur la valeur de la ou des variables qui rendent fausse l'hypothèse — les "hors-sujet" de M. Legrand (1990).

La méthode choisie a donc été de faire passer un test (non anonyme) aux étudiants en cours d'inscription au Capes au mois de septembre, réunis dans un amphithéâtre, pendant une durée de 4 heures. Les items des deux tests (1999 et 2001) étaient très variés, allant de questions concernant en apparence la vie courante à d'autres très mathématiques (on pourra voir en annexe 2 les items analysés ici).

Types d'implications

La deuxième hypothèse faite, pour le dépouillement, est qu'on peut détecter l'éventail des structurations des modes de validation de l'implication par l'étude des conjonctions de réponses à certains types d'implication présents dans plusieurs items.

Ceci nous a amené à établir une typologie de nos diverses implications au moyen de catégories prenant explicitement en compte les contenus mathématiques ou non mathématiques, tant des assertions en jeu que des modes de validation possibles.

Nous avons ainsi distingué :

* les implications "*calculables*", à des degrés variés (et donc à contenu mathématique) : "$H_n \Rightarrow H_{n+1}$", "si 1=2...", "si $(x^2+1) \leq 0$", les deux items à contenu polynomial ;
* les implications "*arbitraires*", correspondant à la définition d'une "règle" (en général non mathématique) :"Wason", "Radford" ;
* les implications de "*contrat social*": "les bonbons de la maîtresse" ;
* les implications "*factuelles*", *non calculables*, où les assertions P et Q (hypothèse et conclusion) sont des données immédiatement saisissables par le sujet, pratiquement sensibles, et qui peuvent éventuellement être à contenus mathématiques "routiniers" pour les sujets : "Triangle", "Circuit", "Labyrinthe".

Les résultats au premier test : existence de profils types, corrélations de ces types avec le succès à certains types d'items implicatifs

Nous avons lors du dépouillement du test proposé à la première population (Rogalski & Rogalski, 2001) classé les sujets selon quatre profils de comportement, avec variantes pour les trois premiers, en croisant les réponses à la validation de trois implications de type factuel (voir annexe 2) à hypothèse fausse (en fait : toujours fausse, dans la mesure où il y a une variable universellement quantifiée, parfois cachée) : le "circuit électrique" de Marc Legrand (Legrand, 1990), le"labyrinthe" d'Evapm étudié par V. Durand-Guerrier (1996), et le"triangle".

Le choix de ces items comme permettant de tester ces profils repose sur leur aspect factuel ou "matériel", leur immédiate intelligibilité par les sujets, l'absence de référence à des connaissances mathématiques non routinisées (autres qu'immédiates pour les sujets considérés), et la non-calculabilité de l'implication : la proposition "conclusion" ne peut être dérivée de l'hypothèse par un calcul, qui permettrait aux étudiants de dérouler un schème d'enchaînement de calculs simples. Aussi bien le sens des propositions que la valeur de vérité de l'hypothèse et de la conclusion sont aisément évaluables par référence aux propriétés des objets en jeu (pour les sujets adultes cultivés de la population concernée).

Les profils obtenus sont les suivants :

* **logique stable** (10 sujets)
  (réponse du type : "*l'implication est vraie" — en général avec l'argument "parce que l'hypothèse est fausse"*— dans les 3 items) ;
* **logique instable** (9 sujets)
  (la même réponse dans 2 items sur 3) ;
* **pertinent stable** (9 sujets)
  (réponse du type : "*l'implication est stupide", "elle n'a pas de sens"*, dans les 3 items);
* **pertinent instable** (14 sujets)
  (la même réponse dans 2 items sur 3) ;
* **non conditionnel stable** (22 sujets)
  (réponse du type : "*l'assertion est fausse car l'hypothèse est fausse"*, dans les 3 items) ;
* **non conditionnel instable** (23 sujets)
  (la même réponse dans 2 items sur 3) ;
* **sans dominante** ou indifféremment"**mixte**" (20 sujets)
  (tous les autres types de réponses ou de distributions de réponse aux 3 items, incluant les non-réponses éventuelles).

Le terme "non conditionnel" est choisi pour rendre compte du fait que ces étudiants ne conçoivent pas la vérité de l'implication comme un lien conditionnel entre les valeurs de vérité des propositions.

Bien sûr, la question qui se pose alors est la pertinence de cette définition des profils : sont-ils un moyen de prédiction statistique au succès ou à l'échec aux autres items? Y a-t-il corrélation avec le succès au Capes ? En ce qui concerne la population du test 1, la réponse à ces deux questions est développée dans (Rogalski & Rogalski, 2001), et reprise dans (Rogalski & Rogalski, 2003), et c'est chaque fois : oui ! Nous renvoyons à ces textes, et aux graphiques donnés en annexe 1.

Confirmation des résultats au deuxième test

Nous nous sommes naturellement demandé si les résultats obtenus avec le premier test étaient relatifs à la population particulière testée, ou s'il y avait stabilité d'une population à une autre (comparable). Nous avons donc recommencé l'opération test en septembre 2001, exactement dans les mêmes conditions, avec un texte reprenant, parmi d'autres, les mêmes items que ceux étudiés dans le premier dépouillement (avec parfois quelques différences de formulation dont nous voulions tester l'effet, en particulier l'introduction de la forme "*si...alors...*", dans certains items). Disons tout de suite que cette modification a été sans effet sur l'item "triangle" ce qui nous a permis de le garder pour la définition des profils, avec les deux autres implications non calculables à prémisse fausse (de formulation inchangée).

Nous avons construit les mêmes catégories avec les mêmes trois items, et aussi étudié la corrélation entre ces catégories et le succès aux mêmes autres items que dans le premier test. La population ayant baissé (moins de candidats au Capes...), nous avons dû opérer des regroupements pour éviter des dispersions qui auraient rendu les résultats peu significatifs ; précisément, nous avons regroupé les catégories "stable" et "instable" de même type en une seule catégorie ; nous travaillons donc à partir d'ici avec quatre catégories ou quatre profils : "Logique", "Pertinent", "Non conditionnel", et "Sans dominante".

La répartition en catégories de la population testée est alors semblable à celle obtenue avec la première population, ainsi que le montre le tableau 1 ci-dessous.

Tableau 1.Distribution —en pourcentage—des étudiants selon leur profil de réponse aux implications "non calculables à prémisse fausse".

PROFIL	TEST 1 (N=107)	TEST 2 (N=71)
LOGIQUE	17,7	21,1
PERTINENT	21,5	23,9
NON CONDITIONNEL	42	39,4
SANS DOMINANTE	18,7	15,5

Traitement de différentes implications

Les distributions des réponses selon les profils sur chacun des trois autres items de raisonnement sont globalement semblable pour les deux tests, à des détails près qui s'expliquent par les variations de formulation, avec cependant quelques différences locales qui soulèvent des questions dont les réponses sont présentées dans (Rogalski & Rogalski, 2003) auquel nous renvoyons le lecteur.

Tableau 2a. Pourcentages de réponses correctes à l'implication calculable à prémisse fausse ($H_n \Rightarrow H_{n+1}$) selon les profils des étudiants, pour chacune des deux populations.

	LOGIQUE	PERTINENCE	NON CONDITIONNEL	SANS DOMINANTE
TEST 1	73,7	56,5	46,7	35
TEST 2	73,3	29,4	28,6	54,5

Tableau 2b. Pourcentages de réponses correctes (selon la logique mathématique) à l'item de contrat social selon les profils des étudiants, pour chacune des deux populations.

	LOGIQUE	PERTINENCE	NON CONDITIONNEL	SANS DOMINANTE
TEST 1	63,2	21,7	26,7	20
TEST2	60	35,3	35,7	18,2

Tableau 2c. Pourcentages de réponses correctes aux items de sélection (Wason & Radford) selon les profils des étudiants, pour chacune des deux populations.

	LOGIQUE	PERTINENCE	NON CONDITIONNEL	SANS DOMINANTE
TEST 1	68,4	52,2	37,8	30
TEST 2	66,7	52,9	60,7	36,4

Étude de points plus spécifiques

Les tests comportaient 20 items dans le premier passage, et 18 dans le second. De nombreuses questions pouvaient ainsi être étudiées, concernant par exemple : l'usage des quantificateurs (en particulier en rapport avec la notion de contre-exemple et dans des items utilisant des polynômes du second degré dépendant de plusieurs paramètres) ; les corrélations éventuelles entre les profils et certains items très mathématiques (différence entre suite numérique non majorée et suite tendant vers $+\infty$, par exemple) ; l'effet de certaines formulations "à l'envers" ("nul n'est P s'il n'est Q", "parmi les rationnels, seuls les décimaux peuvent avoir 2 développements décimaux illimités différents"...).

Dans un premier temps, nous avons étudié quatre questions centrées sur l'implication elle-même : quelle est l'utilisation de la contraposée ? que se passe-t-il quand on attire l'attention des étudiants sur l'existence de "hors-sujet" (au sens de M. Legrand, c'est-à-dire des x pour lesquels l'hypothèse P(x) est fausse)? comment sont traitées des implications calculables à prémisse toujours fausse ? y a-t-il des effets des changements de formulation de l'implication ?

Nous rappelons ici les résultats sur l'utilisation de la contraposée dans les items communs aux deux tests (ci-dessous). Les résultats sur le traitement des implications calculables concernent le seul test 2 (où on avait spécifiquement ajouté des items de cet ordre) ; on peut les trouver dans (Rogalski & Rogalski, 2003). Le traitement des "hors-sujet" est présenté dans la partie II consacrée aux relations entre profils et traitements ou performances mathématiques. L'effet des changements de formulation a été plus particulièrement introduit dans des visées de comparaison avec des études de psychologie : nous les traitons dans la partie III.

Utilisation de la contraposée pour valider une implication

Sur l'ensemble des 8 items identiques ou à formulations voisines figurant dans les deux tests, nous avons comparé l'usage de la contraposée. Les résultats sont les suivants, en pourcentage des étudiants :

Tableau 3. Utilisation de la contraposée dans les deux tests (sur les items communs).

Utilisent la contraposée	0 fois	1 fois	2 fois	3fois	4 fois
Test 1	59 %	27 %	12 %	2 %	--
Test 2	45 %	24 %	24 %	5,6 %	1,4 %

Au total, 41 % des sujets utilisent au moins une fois la contraposée lors du premier test, et 55 % lors du deuxième test. Le fait qu'un étudiant sur deux, en moyenne, utilise au moins une fois la contraposée est un résultat qui contraste avec une remarque figurant dans (Deloustal-Jorrand, 2000) sur l'absence d'utilisation de la contraposée. Mais son étude clinique concernait un petit nombre d'étudiants, et par ailleurs on peut faire l'hypothèse que les implications à absence de lien sémantique ne déclenchent pas l'usage de la contraposée, même (ou surtout ?) chez des étudiants de mathématiques.

Par ailleurs, il faut noter que la quasi-totalité de l'augmentation d'une population à l'autre entre les deux tests est concentrée dans la réponse à l'item "Wason", dont la formulation avait été changée, la nouvelle formulation en 2001 utilisant un "Si...alors...".

Tableau 4. Pourcentage (par item) d'utilisation explicite de la contraposition dans quatre types de situations d'évaluation d'implications dans les deux tests.

Types de tâches d'évaluation d'implication	Test 1 N=107	Test 2 N=71	Ensemble N=178
Implication non calculable (sur 3items)	3,1	3,7	3,3
Implication calculable ($H_n \Rightarrow H_{n+1}$)	7,5	4,2	6,2
Vérification de règle : sélection de cartes (Wason)	7,5	29	*
Vérification de règle : évaluation de procédures	4,7	10	6,8
Évaluation de contre-exemples **	23,4	32,4	27

* Le fort effet, sur ce seul item, de la différence de formulation ne donne pas de sens à un regroupement des populations des deux tests.

** Il s'agissait d'exhiber un contre-exemple à des affirmations d'étudiants concernant des propriétés de fonctions.

II. "Profils" et réponses aux items mathématiques

Pour étudier les corrélations entre les profils et les performances mathématiques, nous allons analyser trois items où les mathématiques interviennent plus directement, et étudier le succès global des étudiants au concours du Capes. Le premier item étudie les réactions des étudiants, dans le contexte d'inégalités concernant des polynômes du second degré, constatant que dans l'étude d'une implication du type « $\forall x \in X\ (P(x) \Rightarrow Q(x))$ » il peut y avoir des "hors-sujets", c'est-à-dire des éléments x de X pour lesquels P(x) est fausse (voir Legrand, 1990). Les deux items suivants sont purement mathématiques : le premier est une tâche de traduction en langage mathématique formel d'une propriété du développement décimal illimité (DDI) d'un nombre rationnel ; le second demande de trouver toutes les implications et non-implications entre 6 propriétés éventuelles d'une suite de nombres réels. Enfin, le succès ou l'échec au Capes est un indicateur très global, qui teste l'ensemble des performances mathématiques sur des problèmes du niveau des deux premières années d'université et des connaissances et savoirs faire à propos des cours et exercices portant sur les trois dernières années de l'enseignement secondaire (lycée) et la première année d'université.

1. Les "hors sujet"

Dans la conduite des situations de débat scientifique, M. Legrand a identifié trois types de cas particuliers ou d'événements élémentaires apparaissant dans le débat : des *exemples* qui vérifient l'hypothèse et la conclusion de l'assertion, des *contre-exemples* qui vérifient l'hypothèse et pas la conclusion, et des "*hors-sujet*", qui ne vérifient pas l'hypothèse.

Dans une implication de type "quel que soit x, si P(x) alors Q(x)", il peut y avoir des situations où P(x) est vrai pour certains des x et faux pour d'autres. M. Legrand a proposé d'appeler ces derniers "hors-sujet" pour pouvoir en parler dans les débats, après avoir constaté qu'environ un étudiant sur trois en Deug les assimilait spontanément à des contre-exemples. Cette tendance à traiter ces cas comme des contre-exemples peut conduire ces étudiants à considérer qu'une implication est fausse "quand P(x) n'est pas vraie" (ce qui revient aussi à considérer comme du calcul propositionnel et avec un traitement erroné une affirmation dont la validité devrait être testée dans le calcul des prédicats —cf. Durand-Guerrier).

Deux items du test 2 proposent des implications dans lesquelles existent de tels "hors-sujet" : ils concernent des polynômes du second degré, avec paramètre.

Dans un cas (Q5) une question, qui suit la présentation d'un calcul qui permet d'amorcer la validation de l'implication, conduit à identifier un ensemble de valeurs de la variable l pour lequel l'hypothèse est fausse (si l'étudiant y répond sans erreur de technique mathématique). La question suivante porte sur la validité de l'implication. C'est un des items les plus mal réussis dans les implications calculables (environ 20 % de réponses correctes), bien qu'il soit d'une forme tout à fait usuelle *a priori* pour un étudiant de mathématique.

L'autre item (Q13) qui propose de commenter des réactions d'étudiants est un peu moins mal réussi, mais pose néanmoins problème à une majorité d'étudiants (seulement 41 % de réponses correctes).

Si on considère les réponses en fonction des profils, on constate les points suivants :

– les profils "non conditionnel" et "sans dominante" correspondent aux taux les plus bas de réussite, et aucun étudiant à profil "non conditionnel" ou "sans dominante" ne répond correctement à la fois aux deux items de "hors-sujet" ;

– les étudiants à profil "logique" répondent à 87 % correctement pour l'item de discussion (n° XIII. 2001), mais un très grand nombre d'entre eux ne répondent pas à la question de l'item V. 2001, montrant l'existence des "hors-sujet" (conduisant à un faible taux de succès : 33 %) ;

- les étudiants à profil "pertinent" se situent à un niveau intermédiaire.

Alors que les étudiants sont tout à fait susceptibles de "suivre" un calcul montrant que Q(x) est vérifié pour les x vérifiant P(x), ils sont déstabilisés quand on attire leur attention sur l'existence des hors-sujet. Cela nous semble confirmer la fragilité pour nombre d'entre eux du traitement d'une implication en tant que relation entre antécédent et conséquent, fragilité se traduisant par un glissement de l'évaluation de la validité de l'implication à celle de la validité de l'hypothèse (ou de l'antécédent).

2. Traduction d'une propriété concernant les développements décimaux illimités

L'un des items d'analyse élémentaire est un test de traduction en langage mathématique formel d'une propriété des développements décimaux illimités (DDI) des nombres rationnels (voir annexe 2).

Analyse *a priori* du test DDI

L'assertion proposée peut se traduire ainsi :

"$\forall x \in \mathbb{Q}$, {x a 2 DDI différents}$\Rightarrow${x est un nombre décimal}". On peut aussi écrire la contraposée :

"$\forall x \in \mathbb{Q}$, {x n'est pas un nombre décimal}$\Rightarrow${x n'a qu'un seul DDI}".

Plusieurs erreurs peuvent se présenter.

(1) L'écriture de *l'implication réciproque* : "$\forall x \in \mathbb{Q}$, {x est un nombre décimal}$\Rightarrow${x a 2 DDI différents}". Bien que cette implication soit vraie, ce n'est pas la traduction formelle de l'implication proposée. L'erreur peut provenir de la connaissance que peuvent avoir les étudiants de l'équivalence des deux propriétés en jeu, ou bien de ce que l'implication est donnée dans la langue naturelle en sens inverse de ce qu'on écrirait en mathématiques. Cette erreur est codée dans la suite CONV.

(2) Les étudiants peuvent avoir du mal à comprendre le terme "seulement", en pensant que sa traduction mathématique est la notion d'équivalence.

(3) Il peut y avoir des difficultés dans la compréhension de la forme langagière " ...seuls ...peuvent avoir...", avec une interprétation stricte : "ce n'est pas le cas pour tous les décimaux". Il n'y a alors plus d'implication du tout, puisqu'on peut répartir dans ce cas les nombres décimaux en deux catégories : ceux qui ont deux DDI, et ceux qui n'en ont qu'un...

Par exemple, parmi les 50 étudiants du test 2 qui ont répondu à la tâche DDI, 8 ont commis les erreurs d'interprétation (2) et (3).

Ces deux erreurs et des variantes sont regroupées dans notre analyse avec les non-réponses, sous le code NOIM (pas d'implication donnée).

Ces trois erreurs révèlent des problème de traduction en langage formel. Le terme "seuls" doit être compris comme "c'est seulement dans le cas où...", formulation typique de la *condition nécessaire* (CN) que signifie une implication. De même, le terme "peuvent avoir" doit être compris comme associé à la *condition suffisante* (CS) exprimée par une implication. Le sens de l'assertion est "c'est seulement s'il est décimal (CN) qu'un nombre rationnel peut avoir deux DDI différents (CS)". Une bonne compréhension de l'assertion donnée en langue naturelle demande une interprétation de la relation entre une CN et une CS.

Corrélations entre profils et succès au test de traduction formelle sur les DDI

Les réponses des 178 étudiants sont présentées dans le tableau 5 ci-dessous.

Tableau 5. Distribution des types de réponses au test DDI selon les profils (effectifs).

Réponses → Profils ↓	CORRECT	CONV	NOIM	Total
Logique	18	9	7	34
Pertinent	19	8	13	40
Non condition.	19	31	23	73
Mixte	11	7	13	31
Total	67	55	56	178

Commentaires

(1) Seulement 38 % des étudiants traduisent correctement l'assertion proposée ; 31 % ne donnent aucune implication, et 31 % traduisent la réciproque.

Cela montre que les étudiants ont, globalement, des difficultés notables dans la traduction "technique" du langage (mathématique) naturel, même dans le cas d'un contenu mathématique simple (pour leur niveau mathématique : étudiants licenciés avec 3 ou 4 ans d'études universitaires).

(2) Les profils "logique" et "pertinent" réussissent assez bien, de façon similaire (53 % et 47,5 % donnent une réponse correcte), en contraste avec les performances assez faibles des étudiants "non-conditionnel" (42 % d'entre eux donnent la réponse CONV et seulement 26 % la réponse correcte) et des étudiants à profil "mixte" (42 % ne donnent pas d'implications, 35 % répondent correctement).

(3) Les réponses correctes au test DDI proviennent principalement des étudiants "logique" et "pertinent" (55 %, alors qu'ils représentent 42 % de l'effectif total). Inversement, la réponse CONV est donnée principalement par les étudiants à profil "non conditionnel" (56 %, alors qu'ils sont 41 % de l'effectif).

En définitive, la capacité à résoudre un problème de traduction formelle semble relativement corrélée aux profils des étudiants, au moins si nous regroupons les catégories "logique" et "pertinent" d'une part, et "non conditionnel" et "sans dominante" de l'autre.

3. Test sur les implications et non implications entre six propriétés d'une suite de nombres

Le deuxième test purement mathématique demande aux étudiants d'énoncer toutes les implications et non implications existant entre six propriétés éventuelles d'une suite de nombres réels (voir l'énoncé à l'annexe 2).

Analyse *a priori* du test sur les suites

Analyse mathématique
Les implications et non implications sont résumées dans le réseau suivant :

$$\textbf{(a)} \Leftrightarrow \textbf{(f)} \underset{\not\Leftarrow}{\Rightarrow} \textbf{(c)} \Leftrightarrow \textbf{(d)} \Leftrightarrow \textbf{(e)} \underset{\not\Leftarrow}{\Rightarrow} \textbf{(b)}$$

Les assertions (a) et (f) signifient que u_n tend vers $+\infty$ (comme (f) le dit explicitement), (b) affirme explicitement que u_n n'est pas bornée, et (c), (d) et (e) sont différentes manières d'écrire que u_n n'est pas majorée. Ainsi, le réseau d'implication peut se réécrire en langue naturelle ainsi :

$$(u_n \text{ tend vers } +\infty) \underset{\not\Leftarrow}{\Rightarrow} (u_n \text{ n'est pas majorée}) \underset{\not\Leftarrow}{\Rightarrow} (u_n \text{ n'est pas bornée})$$

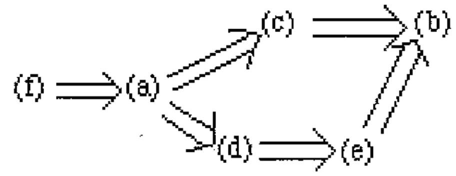
L'équivalence entre les propriétés (a) et (f) est spécifique des suites (il n'y a pas d'analogue pour les fonctions). Elle dépend du fait qu'un sous-ensemble de N (les entiers) est fini si et seulement si son complémentaire contient un intervalle $[A, +\infty[$. Cette équivalence doit utiliser une "traduction" entre un aspect sémantique (en langue naturelle) et une expression formelle d'une propriété.

Une interprétation de (c), (d) et (e) pourrait être : u_n a une sous suite u_{n_p} qui tend vers $+\infty$; une telle forme de ces propriétés pourrait apparaître dans une preuve de l'équivalence entre (c) et (d).

L'équivalence entre (d) et (e) renvoie à la compréhension du rôle de la variable quantifiée r dans l'expression qui signifie qu'une suite n'est pas majorée ($\forall r\ \exists n$ tel que $u_n \geq r$) : r peut être astreint à vérifier une condition supplémentaire qui lui permet d'être "aussi grand qu'on veut". La compréhension de ce point par les étudiants révèlerait la prise de conscience du fait que les inégalités avec quantificateurs qui expriment des concepts d'analyse (être convergente, bornée,...) *ne signifient toujours que des conditions suffisantes*. Ainsi, cette prise de conscience est corrélée avec une conception correcte de l'implication.

Enfin, il faut noter que l'implication (d)$\Rightarrow$(c) est difficile à prouver si on veut construire l'infinité d'entiers demandée (cela renvoie aussi à la construction éventuelle d'une sous-suite...) ; elle est bien plus simple à prouver par contraposition.

Principe d'analyse des réponses



Dans le test sur les suites, on ne demandait pas explicitement aux étudiants de prouver leur réponse. Presque aucun des 178 étudiants n'a donné de justification. La structure très fréquente des réponses est une liste d'assertions : (d) $\Rightarrow$ (b), (c) $\not\Rightarrow$ (f), (d) $\Rightarrow$ (e), etc... Plus rarement, les étudiants écrivent des équivalences telles que (d)$\Leftrightarrow$(e) ou des suites d'implications ou parfois des réseaux d'implications comme sur le dessin de la figure. Dans ces cas, nous avons codé les réponses comme si ces suites ou ces réseaux d'implications contenaient celles dérivables de celles écrites par transitivité (par exemple, dans la figure : (a)$\Rightarrow$(b), (f)$\Rightarrow$(e), etc...).

Il y a ainsi 30 implications ou non implications (en abrégé : I ou NI) correctes possibles, en fait 19 I et 11 NI ; les scores des étudiants sont référés à ces nombres.

Globalement, nous avons utilisé deux types d'analyse.

(1) La première est basée sur le nombre de I et NI correctes explicitement ou implicitement (par transitivité) données dans la réponse, et sur le nombre de I et NI données, qu'elles soient justes ou fausses.

(2) La seconde se focalise sur les I et NI reliant respectivement : (f) et (a) ; (f) et (b) ; (d) et (e). Et cela à cause du contenu mathématique important de ces I et NI. Par exemple, les relations entre (f) et (b) mettent en jeu un aspect important de la sémantique mathématique sur les suites ; celles entre (f) et (a) sont reliées à la formalisation de (f) en terme de quantificateurs et d'inégalités; enfin les relations entre (d) et (e) mobilisent, on l'a déjà dit, la compréhension du fait que les implications standard de l'analyse ne sont que des conditions suffisantes.

Mais dans la présente rédaction nous n'aborderons que les relations entre (f) et (b) et leur aspect sémantique. Nous nous attendions d'ailleurs à un bon score, compte-tenu du niveau d'étude des sujets.

Corrélations entre les profils et les performances globales sur le test sur les suites

Nous signalons d'abord brièvement un résultat un peu inhabituel. Dans les deux années (1999 et 2001), les hommes réussissent mieux que les femmes (plus de I et NI exprimées : 13, 8 en moyenne contre 10,2 ; plus de réponses correctes : 70 % contre 59 %). Ce résultat contredit certains travaux sur la comparaison entre les performances mathématiques des femmes et des hommes. Dans l'étude présentée ici, la différence renvoie sans doute déjà à la différence de la distribution des profils dans la population : 42 % des femmes sont dans le profil "non conditionnel", mais seulement 29 % des hommes (voir Annexe 3). Cependant, cette différence entre les hommes et les femmes est beaucoup moins présente dans les performances au test sur les DDI (la différence apparaît seulement dans le fait que 46 % des femmes à profil "non conditionnel"donnent la réponse CONV, mais seulement 33 % des hommes de ce même profil). Les explications à cette différence hommes-femmes sont sans doute complexes, et demanderaient une étude plus fine de tous les items des tests des deux années, et une étude des raisons qui amènent à la préparation au Capes des étudiants des deux sexes.

Concernant les performances au test sur les suites, le tableau 6 les présente en relation avec les profils. Dans les lignes, nous présentons les nombres moyens de réponses (sur les 30 possibles) dans chaque profil.

Tableau 6. Performances dans le test sur les suites : moyennes des I et NI, exprimées et correctes, en fonction du profil.

	Profil logique	Profil pertinent	Profil non cond.	Profil mixte	Total
I ou NI exprimées	11.6	12.7	11.8	8.7	11.5
I ou NI correctes	8.6	8.7	7	5.8	7.5

Commentaires

Ce test apparaît comme *très difficile* pour les étudiants. Le nombre moyen de I ou NI explicitement données dans les réponses est de 11,5 (pour un maximum de 30). Le nombre moyen de I ou NI corrects est très bas : seulement 7,5. En moyenne, seulement 38 % des I ou NI possibles sont présentes dans les réponses, et 65 % d'entre elles sont résolues correctement, autrement dit il n'y a que 25 % des I ou NI possibles correctement avancées.

La proportion des relations exprimées qui sont correctes est de 83 % pour les implications, et seulement 59 % pour les non implications.

Les profils "logique" et "pertinent" expriment ensemble en moyenne 12,2 I ou NI, avec 8,6 réponses correctes (un score de 70 %), et les profils "non conditionnel" et "mixte" expriment ensemble en moyenne 10,9 I ou NI, avec 6,7 réponses correctes (un score de 61 %). Les profils "logique" et "pertinent" ont des résultats semblables. Il apparaît qu'effectivement, en moyenne, les deux premiers profils donnent un meilleur pronostic pour le succès à ce test que les deux derniers.

Performances des étudiants d'un point de vue sémantique :
I et NI entre les assertions (b) et (f)

La relation entre (b) et (f) a une nature sémantique, car elles expriment des propriétés classiques de suites et sont toutes deux données en langage naturel : (b) est "la suite n'est pas bornée" et (f) est : "la suite tend vers $+\infty$". Regarder les I et NI exprimées par les étudiants devrait nous renseigner sur leurs conceptions sur ces notions.

Nous donnons dans le tableau 7 la distribution des réponses des étudiants.

Tableau 7. Distribution (en pourcentage) des I et NI entre (b) et (f) exprimées, en fonction des profils.

	Logique	Pertinent	Non cond	Mixte	Total
(f)$\Rightarrow$(b) et (b)$\not\Rightarrow$(f) totalement correct	32.4	40	17.8	22.6	26.4
(f)$\Rightarrow$(b) seulement	38.2	17.5	21.9	12.9	22.5
(b)$\Rightarrow$(f) exprimée (fausse)	11.8	15	26	16.1	19.1
aucune relation exprimée	17.6	27.5	34.2	48.4	32

Commentaires

D'abord, il est très surprenant que plus de la moitié des étudiants n'expriment pas l'implication (f)$\Rightarrow$(b), et que presque la moitié d'entre eux donnent l'implication réciproque fausse (b) $\Rightarrow$ (f).

En ce qui concerne la corrélation avec les profils, les étudiants de profils "logique" ou "pertinent" réussissent nettement mieux que les "non conditionnel" ou les "sans dominante" : 63,5 % des deux premiers profils donnent une réponse sans l'erreur (b) $\Rightarrow$ (f), et seulement 38,5 % des deux derniers profils. Globalement, presque 40 % des "non conditionnel" ou"mixte" n'expriment aucune I ni NI, pour seulement 23 % des "logique" ou "pertinent". Il semble que plus de 60 % des étudiants à profil non conditionnel et des étudiants "mixte" n'exercent aucun contrôle sémantique sur les implications possibles, pour une suite, entre "être non bornée" et "tendre vers $+\infty$".

4. Comparaison des performances à l'item (b)/(f) du test sur les suites à celles sur le test sur les DDI

Il est intéressant de comparer les performances à l'item de comparaison de (b) et (f) du test sur les suites à celles du test sur les développements décimaux illimités. Nous présentons les termes de cette comparaison dans le tableau 8.

Tableau 8. Relations exprimées entre (b) et (f) en fonction des réponses au test DDI (effectifs).

Type de réponse au test DDI→ Relation entre (b) et (f) ↓	Correct	CONV	NOIM
(f)$\Rightarrow$(b) et (b)$\not\Rightarrow$(f) totalement correct	24	13	8
(f) $\Rightarrow$ (b) seulement	19	12	7
(b) $\Rightarrow$ (f) exprimée (fausse)	4	14	15
aucune relation exprimée	20	16	26
Total (178)	67	55	56

Commentaires

Bien que les contenus mathématiques soient très différents entre ces deux items, clairement les réponses ne sont pas indépendantes. Moins de 6 % des étudiants donnant la formulation correcte dans le test DDI font une erreur concernant les I ou NI entre (b) et (f). Et presque la moitié des étudiants ne donnant aucune implication dans le test DDI n'en donnent aucune non plus dans l'item (b)/(f).

Ce dernier point confirme l'hypothèse que la question "exprimer une implication ou une non implication" peut n'avoir pas de sens pour nombre d'étudiants, principalement parmi ceux ayant un profil "non conditionnel" ou "mixte". Les seules inférences que ces étudiants tirent d'une implication P$\Rightarrow$Q sont du type : "P est vraie, donc Q est vraie" et/ou "P est fausse, donc Q est fausse". Ainsi ils ne peuvent plus exprimer de relation entre P et Q lorsque la valeur de vérité de P reste ouverte (c'est en rapport avec ce que nous avons vu dans la question des "hors-sujet"). De plus, il semble que la sémantique mathématique leur soit d'une aide limitée.

Une analyse plus fine de la comparaison des deux items fait apparaître deux pôles d'étudiants. Un sous-groupe A est formé de 37 étudiants à profil "logique" ou "pertinent" qui répondent correctement au test DDI ; 86 % d'entre eux expriment au moins la relation (f)$\Rightarrow$(b), sans fausse réponse (la moitié sont totalement correctes). Un sous-groupe B comprend 35 étudiants à profils "non conditionnel" ou "mixte" qui ne donnent aucune implication dans le test DDI ; 51 % d'entre eux n'expriment aucune relation entre (b) et (f), et 22 % proposent l'implication fausse (b) $\Rightarrow$ (f) (moins de10 % d'entre eux écrivent la relation correcte entre (b) et (f)).

5. Lien des profils avec les résultats au CAPES

Tableau 9. Pourcentages d'admissibles et de reçus au Capes selon les profils.

Profils	LOGIQUE	PERTINENCE	NON CONDITIONNEL	SANS DOMINANTE
TEST1				
admissibles	53	52	35	31
reçus	26	30	16	25
TEST2				
admissibles	66,7	52,9	53,6	45,5
reçus	33,3	17,6	28,6	27,3

Globalement, il n'y a pas une différence marquée entre les deux populations pour l'admission (22,5 % pour le test 1 *vs*-26,8% pour le test 2), alors qu'elle est forte pour l'admissibilité, qui est nettement plus importante dans le test 2 (près de 55%). Le profil "logique" est dans les deux cas le plus performant. Il apparaît en revanche des différences pour les autres profils. Le profil "non conditionnel" a de bien meilleures performances dans le test 2. Quant au profil "pertinent", il présente le plus faible pourcentage de reçus.

Il n'est pas évident de savoir à quoi peut être due la forte atténuation, du test 1 au test 2, de la différence entre les profils pour l'admissibilité au CAPES. Une hypothèse, mais que nous n'avons pas testée, serait que la diffusion pendant l'année 2000 des résultats au test 1 et de leur analyse auprès des enseignants du CAPES a pu avoir un effet sur les enseignants, qui, même inconsciemment, seraient plus intervenus sur les questions de logique : cela pourrait expliquer partiellement une certaine atténuation des différences entre profils entre le passage du test 2 en septembre 2001et l'écrit du CAPES en mars 2002.

III. Les approches de la psychologie

Les études sur l'implication sont développées depuis quatre décennies dans des situations expérimentales qui s'intègrent dans le champ plus large des études sur le raisonnement et la rationalité humaines, de Piaget aux études de neuropsychologie sur les processus d'activation / inhibition dans le raisonnement.

1. Qui est rationnel ?

La référence "princeps" est le travail de Wason (1966) visant à mettre à l'épreuve l'existence d'un "stade des opérations formelles" tel que Piaget le postulait. Selon Piaget, ce stade de développement était celui où le sujet (enfant ou adolescent) devenait capable de traiter des propositions indépendamment de leur contenu. (Le stade précédent selon Piaget étant celui des "opérations concrètes" où le raisonnement porte sur des contenus sémantiques, et ne peut en être détaché).

A ce stade, le sujet devient capable de faire des tests d'hypothèses caractéristiques du raisonnement scientifique : la "logique de l'enfant" rejoint la logique scientifique. Pour tester si, conformément au modèle de développement, l'adulte "éduqué" était au stade des opérations formelles, Wason a élaboré deux tests désormais classiques, l'un sur le raisonnement déductif lié au test d'hypothèse : le test de sélection de cartes, l'autre sur le raisonnement inductif lié à l'élaboration d'hypothèse : le test de la complétion de suite de nombres.

Le test de la sélection de cartes a été l'objet d'une multitude de reprises, et de variations : il s'agit dans tous les cas de sélectionner des informations à prendre pour savoir si une règle donnée est vérifiée ou non. Les variations ont concerné la sémantique de la règle (son contenu et sa nature par rapport à l'expérience quotidienne) ou la consigne de la tâche.

Concernant la consigne les principales versions sont les suivantes : on peut demander quelles cartes doivent être retournées (au minimum) (version "*classique*" de sélection) ou proposer au sujet chaque carte en demandant s'il est nécessaire de la tester (version voisine de celle de *Radford*), ou demander quelles sont les cartes critiques (qui permettent de trancher éventuellement) et celles non pertinentes (qui n'apportent pas d'information sur le fait que la règle est suivie ou non) (version de *focalisation*) ; dans certaines expériences la consigne élimine d'abord les cartes qui ne posent pas de problèmes (pour les adultes) dans la version "classique" : P qui est quasiment toujours donnée et non P qui n'est quasiment jamais donnée (version STRA – sélection *avec réduction d'empan* de cartes à sélectionner) ; certaines versions donnent une information sur le fait qu'il faut retourner deux cartes.

Concernant la sémantique, on a différencié des implications avec règle "abstraite", comme celle utilisée par Wason à l'origine, des implications matérielles, des règles "déontiques". Les résultats les plus marquants et les plus constants concernent la facilité de réponse correcte aux règles déontiques (comme celle qui interdit de servir aux moins de 16 ans une boisson alcoolisée) opposées à la dominante de réponses non "logiques" pour les règles abstraites.

Deux grandes théories s'opposent dans un cadre général de psychologie de traitement de l'information : existence d'une logique mentale propositionnelle (on cherche les opérateurs qui rendent compte du traitement des données et de la consigne, avec l'hypothèse d'un module de traitement logique à la Fodor), ou traitement via des modèles mentaux (Johnson-Laird). Les tâches utilisées par les tenants de l'une ou de l'autre ne se sont pas montrées discriminantes. Une proposition (George, 1997) est que les sujets essaient d'abord de se représenter la règle à travers des modèles de réalisation, et en cas d'échec effectuent (ou essayent d'effectuer) un traitement propositionnel. Ces théories rendent par ailleurs mal compte d'une part de l'existence de raisonnement logique dans des situations complexes (dans le raisonnement scientifique de justification par exemple, ou dans le raisonnement mathématique), et d'autre part des effets de changement de domaine ou de contrat.

Un point très clair sur les théories dans le domaine est fait par Mankeltow (2000), vu du côté de la modélisation logique, ou de la modélisation mathématique de théorie des jeux s'agissant de la prise de décision.

Une approche différente — et plus convaincante que celle conduisant au débat cité ci-dessus — est celle de la pragmatique et de la pertinence (Sperber, Politzer). Par exemple, dans les communications humaines on traite la conditionnelle essentiellement dans son sens de double condition "si A alors B sinon pas B" (exemple classique : "si tu ranges ta chambre, tu pourras sortir jouer", si on voulait un "si alors" logique permettant de toute façon de sortir, on dirait simplement "tu pourras sortir", en insistant sur la conclusion toujours vraie et non sur la condition sans effet). Nous évitons de parler d'équivalence, car le rapport ne concerne pas le modèle logique, mais les implicites de la communication. Par ailleurs, il résulte de l'étude des processus d'influence que "accorder" "si P alors Q" conduit à augmenter l'accord sur la validité de P, et accorder Q encore plus. Donc si P est faux, on récusera l'implication, sauf si la conclusion est clairement aussi fausse que P.

Par ailleurs, dans la vie quotidienne, on s'intéresse le plus souvent aux inférences que l'on peut faire, au sens de "quelles données nouvelles je peux tirer des données que j'ai déjà", et plus rarement au traitement de ce qui ne permet pas de tirer une nouvelle donnée, ou aux relations en tant que telles. Par ailleurs, on cherche en général des réponses satisfaisantes plus que des réponses optimales : c'est la théorie de la rationalité limitée, initiée par Simon pour rendre compte du décalage systématique entre les modèles de prise de décision rationnelle et l'activité effective (dominante) des décideurs.

Un développement théorique introduit par Cosmidès est même que le traitement que font "spontanément" de l'implication les adultes, même cultivés, tient à un processus d'adaptation : ce qui a été favorisé par l'évolution - sur le long terme de la phylogenèse - est ce qui est adapté à la vie courante, où il suffit de manière dominante de tirer des conséquences positives et où on ne se règle pas par les exceptions (il suffit qu'une relation soit suffisamment souvent observée pour être traitée efficacement comme une implication). ("*We propose the principle of pre-emptive specificity - that the human cognitive architecture should be designed so that more specialized inference systems pre-empt more general ones whenever the stimuli centrally fit the input conditions of the more specialized system. This principle follows from evolutionary and computational considerations that are common to both relevance theory and the ecological rationality approach*", Fiddick & Cosmides, 2000). Cette dernière théorisation est impuissante à rendre compte du fait qu'il y a autour de 10% des sujets qui répondent de manière logique, dans la population des sujets habituels des expériences (à savoir en général des étudiants en psychologie ou en "*management*"), et qu'un raisonnement scientifique s'est effectivement développé ...

Une extension de l'approche de la pertinence est la prise en compte du contrat dans lequel se situe la tâche (Politzer). Le sujet fait des hypothèses sur ce qui est attendu de lui, en fonction de la situation dans laquelle il se représente être par rapport à celui qui lui soumet une tâche de raisonnement. Il utilise évidemment son expérience pour cela. Cette approche rend compte d'un grand nombre de résultats, y compris ceux où le positionnement qu'on propose au sujet de prendre a un effet sur ses réponses. On retrouve dans ce cadre théorique un élément présent dans les théories de l'activité : l'activité d'un sujet est doublement déterminée et régulée : par les propriétés de la situation et par ses propres propriétés, ici entre autres d'une part sa compétence logique – les instruments cognitifs dont il dispose en ce domaine – et d'autre part son expérience.

Enfin, l'articulation de la psychologie du développement et de la neuropsychologie conduit à mettre en avant les processus d'activation et d'inhibition qui sont en jeu dans les réponses à des tâches comme celles de Wason. Schématiquement, si les expériences dominantes conduisent à des "biais cognitifs" en activant des schèmes non conformes à la logique (comme ceux postulés par la pragmatique ou les schèmes précocement construits dans l'expérience du petit enfant), des processus d'inhibition peuvent les "contrer", et permettre l'expression d'un rationnel construit lui aussi au cours du développement. Cela permet d'expliquer "la coexistence possible du *rationnel construit*, la compétence déductive, et de l'irrationnel présumé révolu : sa transgression par les biais de raisonnement" (Houdé, 1997, p. 27).

Nous retiendrons pour notre part le cadre d'une théorie de l'activité, issue des chercheurs ayant travaillé avec Vygotsky, intégrant dans les déterminants du côté du sujet, à côté de sa compétence dans le champ de la logique, les schèmes d'inférences qui peuvent s'activer ou doivent être inhibés pour répondre aux exigences logiques d'une tâche de raisonnement.

La figure 1 résume la présentation schématique que nous venons de faire.

2. La réponse des scientifiques dans les tests de sélection des cart es

Un certain nombre d'études de psychologie ont proposé la tâche des cartes de Wason à des scientifiques de différents domaines. Diverses formes de la tâche ont été utilisées. Les résultats montrent que les scientifiques dans les conditions « habituelles » de la passation des tâches expérimentales sont meilleurs que les sujets « lambda », c'est-à-dire pour l'essentiel des étudiants de psychologie ou de sciences de la gestion.

Ainsi, Tweney et Yachinin (1985) utilisant la tâche classique de sélection trouvent de meilleures performances pour les scientifiques. Kern, Mirels et Hinshaw (1983) comparent des scientifiques – psychologues, biologistes ou physiciens– sur la tâche des cartes où il s'agit d'identifier les cartes non pertinentes et les cartes critiques (celles qui invalideraient la règle). Les proportions de

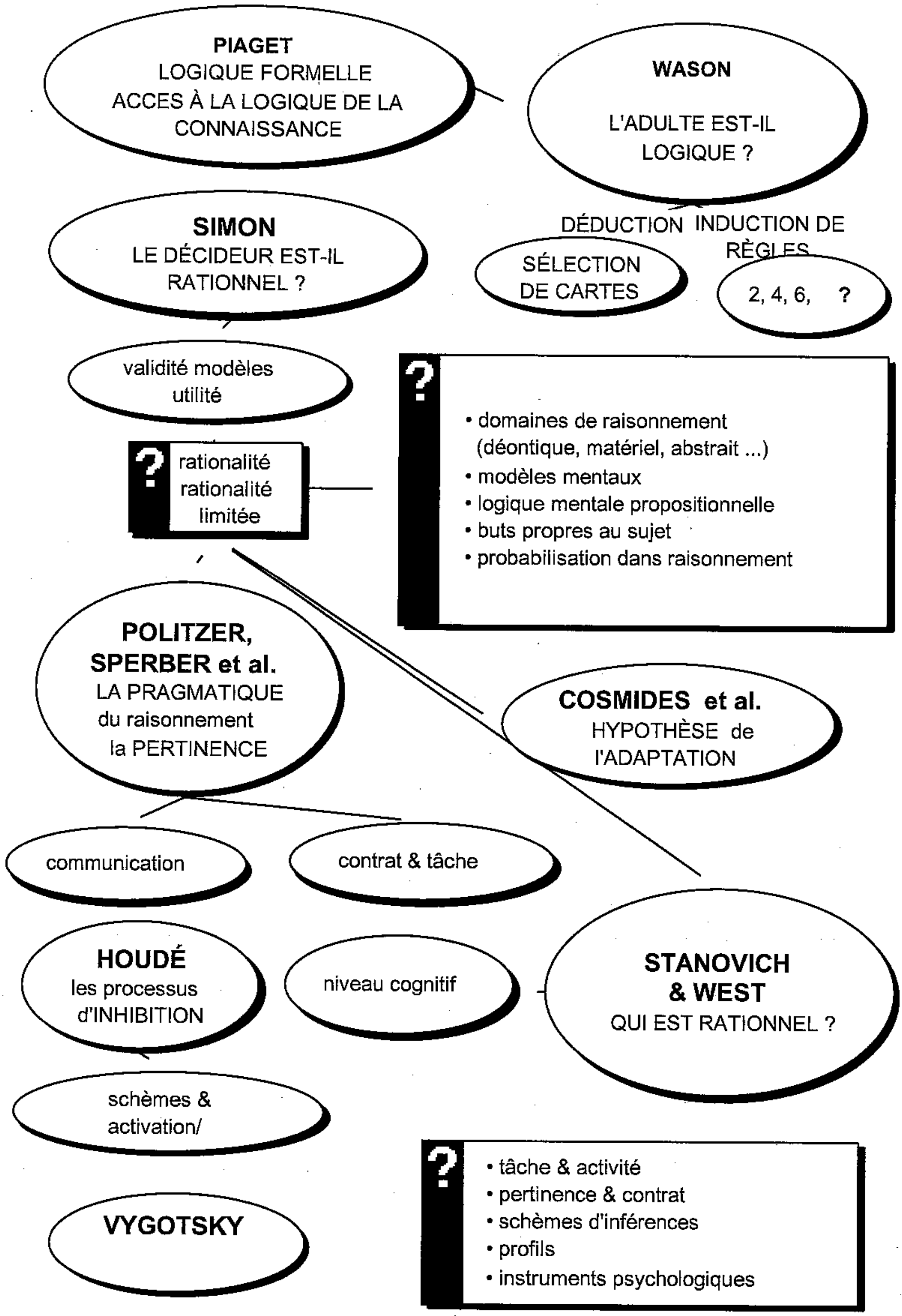


Figure 1. *Schéma du champ des recherches en psychologie cognitive sur le raisonnement et la rationalité humaine.*

réponses correctes (logiques) sont de l'ordre de 15% pour les psychologues, 11% pour les biologistes, et 23% pour les physiciens (il n'y a pas eu d'interrogation de mathématiciens). Quant aux étudiants ayant « suivi un cours de logique » environ 14% d'entre eux répondent correctement. C'est la même proportion que celle obtenue par Love et Kessler (1995) avec des étudiants « lambda », ce qui indique que la focalisation sur les cartes pertinentes améliore légèrement les réponses par rapport à la tâche de sélection classique.

Par ailleurs, dans une recherche sur "*Qui est rationnel*", une étude différentielle mettant en relation les scores aux tests d'entrée en université (dits scores SAT) montre que ces scores SAT sont meilleurs pour les étudiants réussissant la tâche classique de sélection de cartes (Stanovich, & West, 1998).

Les conclusions tirées par les auteurs de ces études sur la rationalité des scientifiques ont été discutées par Politzer, qui a mis en évidence un fort effet de contrat dans cette tâche de Wason. Il a en effet présenté la tâche classique de sélection à des élèves de Polytechnique (issus donc de classe préparatoires scientifiques à dominante à l'époque mathématique) dans deux situations différant par le contrat : dans la première situation, la tâche a été présentée comme une tâche expérimentale par un psychologue, dans la seconde situation, elle a été présentée explicitement comme une tâche sur l'implication. Les réponses correctes passent alors de 25% dans la première situation (celle des études habituelles des psychologues) à 60% dans la seconde situation en appelant au champ conceptuel de la logique (déductive) (Politzer, séminaire de laboratoire, juin 2001).

Les résultats obtenus dans nos tests vont dans le sens des interprétations proposées par Politzer, et rejoignent même précisément les valeurs qu'il avait obtenues avec des étudiants scientifiques sélectionnés, dans la condition d'un contrat en appelant au champ de la logique— ce qui était le cas avec les étudiants de nos tests.

Les études de psychologie cognitive se sont centrées sur l'explication des 90% de réponses non "logiques", et la prise en compte du contrat et de la compétence logique formée dans l'activité d'étudiant en mathématiques rend compte du taux substantiel de réponses logiques chez nos étudiants. Mais, la psychologie cognitive reste discrète sur les 10% de réponses correctes chez les sujets "lambda" et le problème demeure des 40% de réponses non conformes à la logique mathématiques chez des étudiants futurs professeurs de mathématiques ...

3. L'importance des formulations ? Réponses aux implications calcula bles selon l'usage ou non du "si ... alors...".

Les changements de formulation ont concerné trois des items d'implications. Lors du premier test, des formulations "non canoniques" avaient été utilisées, et la forme retenue sans utilisation du "si" conditionnel est inhabituelle dans les études de psychologie : pour comparer notre population de futurs enseignants de mathématiques aux populations de ces travaux nous avions besoin de reprendre leur formulation. Par ailleurs, l'utilisation de formes où le conditionnel de surface n'était pas apparent pouvait modifier les traitements par les étudiants (Politzer, communication privée).

La formulation non canonique pouvait avoir trois types d'effets : 1) "bloquer" le recours au traitement de la conditionnelle dans la vie quotidienne (moins d'effet de "pragmatique") ; 2) empêcher au contraire des étudiants de reconnaître l'implication (conduisant à plus d'erreurs), par exemple la formulation avec le terme "tout" ("tout triangle" ou "toutes les boules blanches") pouvaient cantonner des sujets à ne se placer que dans le cas de l'hypothèse vraie, favorisant ainsi a priori un mode d'évaluation de type "pertinent" ; 3) ne pas avoir d'influence sur les réponses, étant donné le niveau des étudiants passant le test (ce qui était notre propre hypothèse).

Nous avons donc effectué les modifications de formulation pour utiliser la formulation canonique "*si ... alors ...*" dans les tâches de sélection (version Wason et version Radford), et les trois implications non calculables à prémisse fausse (Triangle, Circuit et Labyrinthe). Nous donnons ci-dessous, dans la figure 2, les versions utilisées respectivement dans Test 1 et Test 2.

On observe effectivement un changement dans la réussite à la tâche classique de sélection de

Wason : la formulation « si ..alors .. » s'accompagne d'une augmentation de la proportion de réponses « logiques » d'un peu moins de 48% à près de 65%. (Khi2 (dl=1,N=178) = 5, p<.02). On a vu que la formulation en « si …alors… » s'accompagne d'une augmentation de l'utilisation de la contraposition pour le profil « non conditionnel ». Néanmoins, l'effet n'est pas très important et la non-prise en compte du nombre impair reste l'erreur dominante, avec plus de 30 % des réponses.
En revanche, sur l'item "Radford" il n'y a pas d'effet du changement de formulation. Par ailleurs, l'impact sur les réponses correctes à la fois à la tâche classique (« Wason ») et à celle de l'évaluation de procédures (« Radford ») est en fait très modéré (on passe de 45% dans le test 1 à un peu plus de 56% dans le test 2). En fait, le pourcentage d'étudiants répondant correctement à « Radford » après avoir répondu correctement à « Wason » passe de 94% dans Test 1 à 87% dans Test 2. De fait, avec l'usage de la formulation conditionnelle "canonique" "*Si... alors ...*", les distributions de réponses aux deux versions de la tâche de sélection de cartes "Wason" et "Radford" se rapprochent, ce qui ne va pas dans le sens des résultats obtenus dans les expériences de psychologie avec des populations "tout venant".

Enfin, le changement de formulation a un effet tout à fait négligeable sur l'item "Triangle" (un peu plus de réponses « vrai » et un peu moins de réponses « non pertinent », mais pas statistiquement significatif, et pas de changement sur les réponses « assertion fausse » ou « autres »).

Pour interpréter cette très minime modification, il faut la comparer à celle tenant simplement à la variation de population. Le tableau ci-dessous montre qu'il apparaît la même tendance à un peu plus de réponses logiques également pour les deux autres implications, bien que leur formulation soit inchangée : on reste dans les variations statistiques attendues sous l'hypothèse d'une absence d'effet de formulation, et des populations globalement semblables.

<table>
<tr><th>versions du test 1</th><th>versions du test 2</th></tr>
<tr><td colspan="2">• Wason sélection de cartes
la règle suivante</td></tr>
<tr><td>"derrière une voyelle il y a un chiffre pair".</td><td>si une carte a une voyelle sur une face, alors elle a un nombre pair sur son autre face".</td></tr>
<tr><td colspan="2">• Radford choix de procédures
on s'intéresse à</td></tr>
<tr><td>la question suivante : "est-ce que, dans l'urne, toutes les boules blanches ont un numéro pair?".</td><td>la véracité de l'assertion suivante: "dans l'urne, si une boule est blanche, alors son numéro est pair"</td></tr>
<tr><td colspan="2">• Implication factuelle à prémisse fausse (Triangle)
que pensez-vous de l'assertion suivante ?</td></tr>
<tr><td>"Tout triangle non aplati du plan, dont les médiatrices ne sont pas concourantes, est équilatéral".</td><td>"Si un triangle non aplati du plan a ses médiatrices non concourantes, alors il est équilatéral".</td></tr>
<tr><td colspan="2">• Implication factuelle à prémisse fausse (Circuit de Legrand) (inchangée)
Que peut-on dire de la vérité de l'assertion suivante</td></tr>
<tr><td colspan="2">"si L1 est allumée et si L3 est allumée, alors L2 est allumée et L5 n'est pas allumée"</td></tr>
<tr><td colspan="2">• Implication factuelle à prémisse fausse (Labyrinthe de Durand-Guerrier) (inchangée)
dire si la phrase est vraie, si elle est fausse ou si on ne peut pas savoir</td></tr>
<tr><td colspan="2">(7) si X est passé par S, il est passé par T</td></tr>
</table>

Figure 2. *Les versions des items Wason, Radford, et d'implications factuelles à premisses fausses.*

Tableau 10. Distribution en pourcentage des catégories de réponses aux implications non calculables à prémisse fausse selon la formulation test1 / test 2

	VRAI	NON PERTINENT	FAUX	X OU RIEN
TRIANGLE				
Test1	14	35,5	43	7,5
Test2	19,5	29,5	45	5,5
CIRCUIT				
Test1	12	31	48	9
Test2	19,5	29,5	35,5	15,5
LABYRINTHE				
Test1	27	22,5	45	5,5
Test2	34	20	41	5,5

Rappel : dans le cas de la tâche du « Labyrinthe », la réponse vraie peut être obtenue non pas dans un raisonnement logique, parce que la prémisse a été évaluée comme fausse, mais du fait d'une non-prise en compte de la condition « ne pas repasser deux fois par la même case », qui permet de rendre vraie la prémisse. Impossible de mettre ceci en évidence dans notre situation de réponse écrite collective.

Conclusion

La compétence logique d'étudiants se préparant à devenir enseignants de mathématiques a été étudiée dans un test comportant un ensemble de questions portant spécifiquement sur l'implication, avec des contenus variés. Deux épreuves d'évaluation de règle (Wason et Radford) ont été introduites pour situer les productions de ces étudiants par rapport à celles étudiées dans des tâches expérimentales de la psychologie. Une attention particulière a été donnée à des situations non étudiées par la psychologie : l'évaluation d'implications à prémisses toujours fausses, ainsi qu'à des "hors-sujets" rendant la prémisse de l'implication fausse.

On a défini une typologie d'implications, susceptibles de provoquer des traitements différents. Trois implications non calculables à prémisse fausse ont servi de base pour définir des "profils" définissant des modes de traitement de ce type d'implication. Un profil "logique" correspond au mode de traitement attendu de mathématiciens : il est le fait d'environ 20% des étudiants. Un profil "pertinent" correspond à un mode de traitement qui, de manière dominante, ne considère une implication que lorsque la prémisse est vraie (et qu'on peut donc en déduire la véracité de la conclusion) ; il regroupe environ 23% des étudiants. Un profil "non conditionnel" traite les implications comme fausses dès que la prémisse est fausse, ne distinguant pas la valeur de vérité de la prémisse de celle de l'implication : 40% des étudiants traitent ainsi les implications à prémisse fausse. Les modes de traitement mixte (variant pour toutes les implications) sont regroupés dans un "profil" sans dominante".

Ces profils sont corrélés aux réponses que les étudiants donnent quand ils doivent tester ou évaluer d'autres types d'implication. Dans leur majorité (de 60% à 70%) les profils "logique" traitent de façon logique ces autres implications ; Les profils "pertinent" sont presque aussi souvent logiques sur les implications calculables et le test d'hypothèse de Wason (50%), mais traitent de manière dominante une implication de contrat social en rendant vraie la prémisse (en interprétant la situation présentée), donc en cohérence avec leur approche de "pertinence". Les autres profils traitent beaucoup moins les autres implications de manière logique. Toutefois, l'utilisation de la contraposée, dans la tâche de test d'hypothèse de Wason permet à certains d'entre eux de retrouver une prémisse positive et de donner en conséquence une réponse logique.

On a ensuite analysé les relations entre ces profils à base logique et le comportement des étudiants confrontés à des tâches de raisonnement en analyse. Une tâche de "traduction" demandait de mettre sous la forme d'une implication quantifiée une relation exprimée en "langage mathématique naturel" (IDE). La capacité à résoudre une telle tâche est apparue liée aux profils des étudiants : les profils "logique" et "pertinent" réussissant de manière analogue (mais seulement pour la moitié d'entre eux), les profils "non-conditionnel" donnant la réciproque comme réponse dominante, alors qu'un grand nombre des profils "mixte" ne proposaient aucune implication.

Une seconde tâche d'analyse demandait d'écrire toutes les implications et les non-implications reliant six assertions sur une suite réelle (Suite). Les réponses étaient corrélées aux profils d'une manière similaire : ici aussi, les profils "logique" et"pertinent" ont exprimé plus de relations, avec moins d'erreurs. Toutefois il s'agit d'une tâche difficile, impliquant trois composants : la compétence logique, telle qu'on peut l'apprécier à travers les profils, la disponibilité des connaissances mathématiques (en particulier pour exprimer des non implications), et la capacité à élaborer une stratégie pour organiser les réponses (nécessaire, vue la combinatoire *a priori*).

Une analyse particulière des implications entre deux propositions exprimant des propriétés simples de suites réelles en "langage naturel mathématique" ont confirmé la corrélation avec les profils : plus de 60% des profils "non-conditionnel" et "mixte" n'exercent aucun contrôle mathématique sur les implications entre les propositions concernées.

La mise en relation des réponses aux deux tâches (IDE et Suite) montre des relations fortes entre les réponses des étudiants, alors que ces tâches diffèrent à la fois dans leur contenu mathématique et dans le type de réponse à produire : les réponses sont meilleures pour les profils "logique" et "pertinent" ; à peu près la moitié des étudiants qui ne donnent pas d'implication pour la tâche IDE ne peuvent pas non plus produire d'implication entre les deux propositions "sémantique" concernant une suite. Les données indiquent que pour nombre d'étudiants parmi les profils "non conditionnel" ou "mixte" la question d'exprimer une implication (et *a fortiori* une non implication) peut manquer de sens. Ces étudiants peuvent faire des inférences d'une implication P=>Q, sous la forme (correcte) comme "P est vraie, donc Q est vraie"ou sous la forme (incorrecte) "P est fausse, donc Q est fausse", mais ils ne peuvent pas exprimer la relation entre P et Q quand leur valeur de vérité reste ouverte. De plus, la sémantique mathématique ne leur semble que d'un secours limité.

Les relations des profils avec le traitement de tâches mathématiques ne se retrouvent pas aussi nettement avec les résultats au CAPES : pour les étudiants du premier test, l'admissibilité au CAPES était nettement plus importante pour les profils "logique" et "pertinent", et plus particulièrement pour les profils stables (même type de réponse aux trois implications non calculables) ; on ne retrouve pas une telle différence pour les étudiants du second test, même si les profils logiques sont un peu plus souvent parmi les reçus (mais de manière générale le taux de réussite a été plus élevé cette année-là).

Notre étude se situe d'une manière originale par rapport à celles de psychologie sur le raisonnement déductif, tout d'abord par le rôle donné aux implications à prémisse fausse et l'approche de la quantification à travers le traitement des "hors-sujets", et ensuite par la définition de types d'implication au-delà de leur domaine sémantique, en distinguant des implications non calculables et des implications calculables, et par la définition de profil dont nous considérons qu'ils reflètent des schèmes de traitement de l'implication. Dans cette approche différentielle on cherche à analyser les différents types de réponses en cherchant la cohérence qui existe entre elles.

L'insertion de tâches reprenant celles d'expériences classiques permet par ailleurs de situer les résultats des étudiants par rapport à ceux des expériences : ils traitent beaucoup mieux les questions d'évaluation de règles abstraites comme celle de la tâche de "Wason", même quand on compare aux — rares—scientifiques réalisant ce type de tâche. L'utilisation de la forme canonique de l'implication en "si ... alors ..." semble déclencher chez certains profils "non conditionnels" une utilisation de l'instrument cognitif que constitue la contraposition, conduisant à un traitement logique, utilisation par ailleurs au moins épisodique chez plus de la moitié des étudiants, ce qui pourrait constituer un point d'ancrage pour développer leur compétence au traitement logique.

De nombreux autres résultats de nos tests restent à analyser pour affiner l'image des schèmes et les représentations de l'implication. Des entretiens conduits avec des étudiants seront également analysés pour une compréhension plus approfondie.

Les résultats conformes à la logique restent cependant d'un niveau loin de ce qu'on pourrait attendre raisonnablement de futurs enseignants de mathématiques (même si un profil "pertinent" permet de faire la plupart des mathématiques suffisantes pour réussir), on peut s'interroger sur la capacité à bien identifier les problèmes que rencontreront les élèves quant à la justification et la preuve mathématique. Nous avons abordé la question de la formation des futurs enseignants (et en amont des étudiants) ; nous n'y revenons pas ici, sinon pour souligner la nécessité pour les enseignants à l'université d'être conscient de la maîtrise insuffisante dont soufrent de trop nombreux étudiants concernant l'implication en général et en mathématique en particulier (même ceux qui n'ont pas été particulièrement en échec et s'intéressent assez aux mathématiques pour vouloir les enseigner).

Références

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

Benbachir, A., & Zaki, M. (2001). Production d'exemples et de contre-exemples en analyse : étude en première année d'université. *Educational Studies in Mathematics, 47(*3), 273-295.

Boero, P. (2000). Entrer dans la culture des théorèmes à 12-14 ans: un défi pour la didactique des mathématiques. In T. Assude & B.Grugeon (Éds.), *Actes du séminaire national de didactique des mathématiques* (pp. 41-54). Paris : DIDIREM Université Paris7.

Deloustal-Jorrand, V. (2000). L'implication. Quelques aspects dans les manuels et points de vue d'élèves-professeurs. *Petit x, 55*,35-70.

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

Durand-Guerrier, V. (1996). Conférence à l'Université de Cornell, USA.

Durand-Guerrier, V. (1996). *Logique et raisonnement mathématique.* Thèse de Doctorat. Université Claude Bernard Lyon 1.

Durand-Guerrier, V. (1999). L'élève, le professeur et le labyrinthe. *Petit x, 50*, 57-79.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics, 53*(1), 5-34.

Durand-Guerrier, V., Le Berre, M., Pontille, M. C., & Reynaud-Feurly, J.(2000). *Le statut logique des énoncés dans la classe de mathématiques. Éléments d'analyse pour les enseignants.* IREM de Lyon.

Duval, R. (1991). Structure du raisonnement déductif et apprentissage de la démonstration. *Educational Studies in Mathematics, 22*, 233-261.

Hanna, G. (2000). Proof, explanation and exploration: an overview. *Educational Studies in Mathematics, 44* (1/2), 5-23.

Hanna, G., & Jahnke, H.E. (1993). Proof and application. *Educational Studies in Mathematics, 24*(4), 421-438.

Houdé, O. (1997). Développement cognitif et inhibition. De l'erreur A-non-B aux biais de raisonnement. *Psychologie Française, 42(*1), 23-29.

Housman, D., & Porter, M. (2003). Proof schemes and learning stratégies of above-average mathematical students. *Educational Studies inMathematics, 53*(2), 139-158.

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

Jones, K., Gutiérrez, A., & Mariotti, M. A. (Eds). (2000). (Special issue on proof in dynamic geometry environments). *Educational Studiesin Mathematics, 44* (1/2).

Kern, L. H., Mirels, H. L., & Hinshaw, V. G. (1983). Scientists' understanding of propositional logic : an experimental investigation. *Social Studies of Science, 13*, 131-146.

Knuth, E. J. (2002). Secondary school teachers' conceptions of proof. *Journal of Research in Mathematics Education, 33*(5),379-405.

Küchemann, D., & Hoyles, C.(2002). Students' understanding of a logical implication and its converse. In A.D. Cockburn & E. Nardi (Eds.), *PME26. Proceedings of the 26th Annual Conference IGPME*, (pp. 3-241–3-248). Norwich: SEPD UEA Norwich.

Legrand, M. (1989). Genèse et étude sommaire d'une situationco-didactique : le débat scientifique en situation d'enseignement, in C. Laborde (Éd.), *Premier Colloque Franco-Allemandde Didactique des Mathématiques etde l'Informatique*. Grenoble : La Pensée Sauvage.

Legrand, M. (1990). "Circuit" ou les règles du débat mathématique, in Commission Inter-Irem Université (CI2U) (Eds.), *Enseigner autrement les mathématiques en Deug A première année*, (pp.129-161). IREM de Paris 7:

Love, R. E., & Kessler, C.M. (1995). Focusing in Wason's selection task : content and instruction effects. *Thinking and Reasoning, 1*(2), 153-182.

Manktelow, K. (2000). *Reasoning and thinking.* Hove: Psychology Press (1ère édition: Taylor& Francis).

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27*(3), 249-266.

Pellissier, R. (2002). L'enseignement du raisonnement scientifique, in *Une expérience d'enseignement à Nice*, publication de la Commission Inter-IREM Université (CI2U), IREM de Lyon

Politzer, G. (2001). *Communication Séminaire Cognition & Activités Finalisées*, juin2001, Université Paris8.

Politzer, G. (2002). Premise interpretation in conditional reasoning. In D.Hartman & L. Macchi (Eds.), *Reasoning and decision making: A hanbook.* Chichester: Wiley.

Radford, L. (1985). *Interprétation d'énoncés implicatifs et traitements logiques. Contribution à la faisabilité d'un enseignement de la logique au lycée.* Thèse de l'Université Louis Pasteur Strasbourg.

Raffalli, C., & David, R. (2002). Apprentissage du raisonnement assisté par ordinateur. *Gazette des mathématiciens, 92*, 48-56.

Recio, A. M. (2001). Institutional and personal meanings of mathematical proofs. *Educational Studies in Mathematics, 48*(1),83-99.

Rogalski, J., & Rogalski, M. (2001). How do graduate mathematics students evaluate assertions with a false premise? In M. van der Heuvel-Panhuisen(Ed.), *Proceedings of the 25th Conference PME25@NL* (pp. 4.113-4.121).Utrecht, NL: Freudenthal Institute Utrecht University.

Rogalski, J., & Rogalski, M. (2003). Contribution à l'étude des modes de traitement de la validité de l'implication par de futurs enseignants de mathématiques. *Annales de Didactique et de*

Sciences Cognitives, 9, 2003,Strasbourg.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57*, 31-95.

Sperber, D., & Wilson, D. (1989/1986). *La pertinence. Communication et Cognition*. Paris : Les Editions de Minuit. (Première édition, 1986, Relevance: communication and cognition. London: Blackwell.)

Stanovich, K.E. (1999). *Who is rational ? Studies of individual differences in reasoning*. Mahwah, NJ, London : Lawrence Erlbaum Associates.

Stanovich, K. E., & West, R. F.(1998). Cognitive ability and variation in selection task performance. *Thinking and Reasoning*, 4(3), 193-230.

Stanovich, K.E., & West, R. F.(2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Science, 23*, 645-726.

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

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

Annexe 1 : relation entre profils et admissibilité au Capes, et entre profils et réponses correctes aux différents types d'implication.

Figure A.

Pourcentage des étudiants admissibles au Capes en fonction de leur profil.

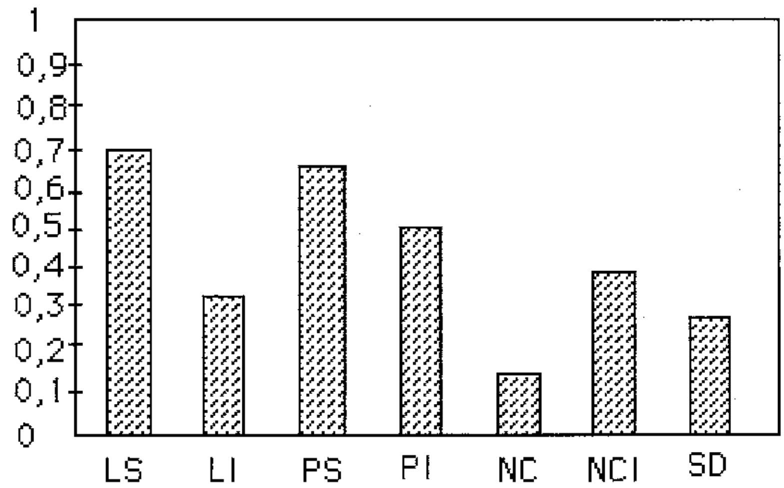
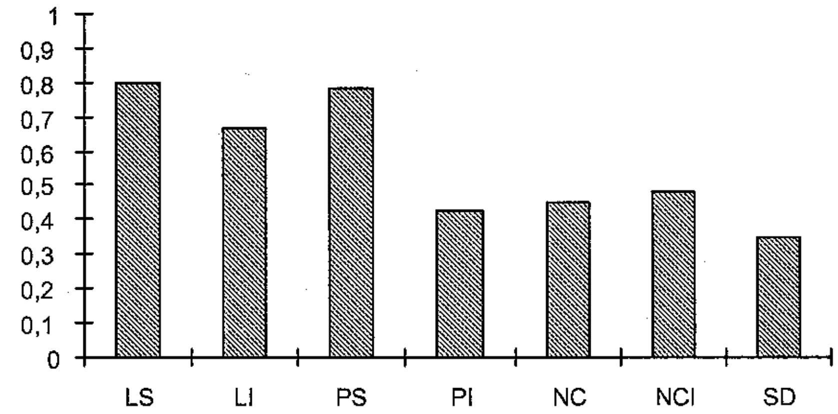


Figure B

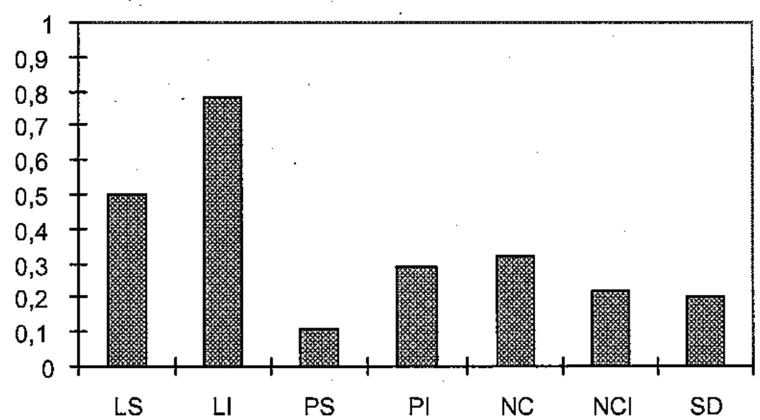
Pourcentage d'étudiants donnant une réponse correcte à l'*implication calculable* $H_n \Rightarrow H_{n+1}$, selon leur profil.



28

Figure C.

Pourcentage d'étudiants donnant une réponse selon la logique à l'*implication de contrat social*, selon leur profil.



Annexe 2 : Texte des items des deux questionnaires (1999, 2001) analysés ci-dessus.

On donne les numéros des items dans chaque questionnaire, en soulignant dans certains cas ce qui est spécifique de chacun des deux tests.

Avertissement (commun aux deux passations): *les phrases, assertions, énoncés, ...mathématiques qui figurent dans les exercices suivants sont souvent donnés sous forme naïve, non formalisée, comme on le fait dans un texte mathématique courant, voire dans la conversation de tous les jours. C'est volontairement, et vous devez vous sentir très libre dans vos réponses, qui peuvent (doivent!) comporter beaucoup de commentaires, et même si vous le voulez aller jusqu'à : "cette question est idiote !".*

Assertion calculable à prémisse fausse

I. 1999 / 2001

Etant donné $l \in R$, on considère la suite $(u_n)_n$ définie par :$u_0 = l$, $u_{n+1} = 2u_n + 1$.

Soit H_n l'assertion " $u_n \leq 2^n / 3 - 1$".

(a) L'implication " $H_n \Rightarrow H_{n+1}$" est-elle vraie pour certains n ? pour tout n ?

(b) Calculer explicitement u_n en fonction de l et de n [on pourra poser $u_n = v_n - 1$].

(c) Montrer que si $l > -2\backslash 3$, toutes les assertions H_n sont fausses.

(d) Si $l = 10$, que peut-on dire de l'assertion " $\forall n\ H_n \Rightarrow H_{n+1}$ " ?

(e) Soit $P_n(l)$ l'assertion "$u_n = 2^{n+5}(l + 1) - 1$". Les assertions suivantes sont-elles vraies ou fausses ?

(1) " $\forall n\ \forall l\ P_n(l)$ est vraie" ;

(2) " $\exists\ l$ tel que $\forall n\ P_n(l)$ est vraie" ;

(3) " $\forall n\ \forall l\ P_n(l) \Rightarrow P_{n+1}(l)$".

Assertion non calculable à prémisse fausse (triangle)

III. 1999

Que pensez-vous de la véracité de l'assertion suivante ?

<u>"Tout triangle non aplati du plan, dont les médiatrices ne sont pas concourantes, est équilatéral".</u>

III. 2001

Que pensez-vous de la véracité de l'assertion suivante ?

<u>"Si un triangle non aplati du plan a ses médiatrices non concourantes, alors il est équilatéral".</u>

Hors-sujet

V. 2001

On se propose d'évaluer la véracité de l'assertion (A) suivante :

(A) : "$\forall x \in R\ \forall l \in R$, si $x^2 - 2lx + 2l + 3 \leq 0$, alors $|x| \leq 2|l| + 1$".

(a) Montrer que l'hypothèse se réécrit sous la forme : $(x - l)^2 \leq (l - 1)^2 - 4$.

(b) Que se passe-t-il si $-1 < l < 3$?

(c) Selon vous, l'assertion (A) est-elle vraie ?

Tâche de sélection de Wason (version classique)

IX. 1999

On dispose de cartes sur lesquelles figurent sur une face une lettre et sur l'autre un chiffre.
On veut tester la règle éventuelle : "derrière une voyelle il y a un chiffre pair". Pour cela on prend un échantillon de 4 cartes, qu'on pose sur la table ; on voit alors la disposition ci-contre .

$\boxed{E}$ $\boxed{4}$ $\boxed{K}$ $\boxed{7}$

Quelle(s) carte(s) exactement doit-on retourner pour savoir si la règle est respectée sur l'échantillon?

VII. 2001

On dispose de cartes sur lesquelles figurent sur une face une lettre et sur l'autre un nombre.
On veut tester la règle éventuelle : "si une carte a une voyelle sur une face, alors elle a un nombre pair sur son autre face". Pour cela on prend un échantillon de 4 cartes, qu'on pose sur la table ; on voit alors la disposition ci-contre .

$\boxed{E}$ $\boxed{4}$ $\boxed{K}$ $\boxed{7}$

Quelle(s) carte(s) exactement doit-on retourner pour savoir si la règleest respectée sur l'échantillon?

Assertion calculable à prémisse fausse (arithmétique)

XI. 2001

L'assertion "si 1 = 2, alors 2 = 3" est-elle vraie ?

Version "Radford" de la tâche de sélection de Wason

XV. 1999

Dans une urne, on dispose d'un certain nombre de boules numérotées. Les boules sont soit blanches soit noires. On s'intéresse à la question suivante : "est-ce que, dans l'urne, toutes les boules blanches ont un numéro pair ?".

On envisage 4 procédures pour répondre à la question :
procédure 1 : on sort de l'urne les boules de numéro pair, puis on regarde leurs couleurs ;
procédure 2 : on sort de l'urne les boules de numéro impair, puis on regarde leurs couleurs ;
procédure 3 : on sort de l'urne les boules blanches, puis on regarde leurs numéros ;
procédure 4 : on sort de l'urne les boules noires, puis on regarde leurs numéros.

Pour chacune des 4 procédures, choisissez parmi les deux options :
(a) la procédure me permettra sûrement de répondre à la question ;
(b) la procédure risque de ne pas me permettre de conclure.

X. Le test sur les DDI

On considère l'assertion suivante :
"Parmi les nombres rationnels, seuls les nombres décimaux peuvent avoir deux développements illimités différents",
(où on précise que par exemple 3, 24 a pour développement décimal illimité 3, 2400000...).
Ecrire cette assertion sous la forme "$\forall x \in X$ (P(x)=>Q(x))", en précisant X, P et Q.

XII. Le test sur les suites

Déterminer toutes les implications et non-implications entre les assertions suivantes concernant une suite réelle arbitraire $(u_n)_n$:

(a) [1999] étant donné un nombre a>0, il n'y a qu'un nombre fini d'entiers n tels que $u_n \leq a$;
[2001] chaque fois qu'on se donne un nombre a > 0, il n'y a qu'un nombre fini d'entiers n tels que $u_n \leq a$;

(b) la suite $(u_n)_n$ n'est pas bornée ;

(c) pour tout M réel, il y a une infinité d'entiers n tels que $u_n \geq M$;

(d) pour tout $r > 0$, on peut trouver un entier n tel que $u_n \geq r$;

(e) pour tout $b > 0$, on peut trouver un entier n tel que $u_n \geq b^2 - b$;

(f) u_n tend vers $+\infty$ quand n tend vers $+\infty$.

Hors-sujet

XIII. 2001

On a posé la question suivante à un étudiant : la proposition (P) suivante est-elle vraie ?

(P) : "$\forall(x, y)\in R^2$, si $y^2 - 2xy + 8x + 9 \leq 0$, alors $|y| \leq 2|x| + 4$".

Il a donné la réponse suivante :

"L'assertion (P) est fausse. En effet, l'hypothèse signifie que l'on a $(y - x)^2 \leq (x - 4)^2 - 25$; mais si $-1 < x < 9$,alors $(x - 4)^2 - 25 < 0$, donc l'hypothèse n'est pas vérifiée, et n'impose donc aucune contrainte à y (on est dans la bande]-1,9[") ; mais la conclusion, dans ce cas, entraînerait sur y la contrainte $|y| \leq 22$, ce qui serait absurde".

Que diriez-vous à cet étudiant ?

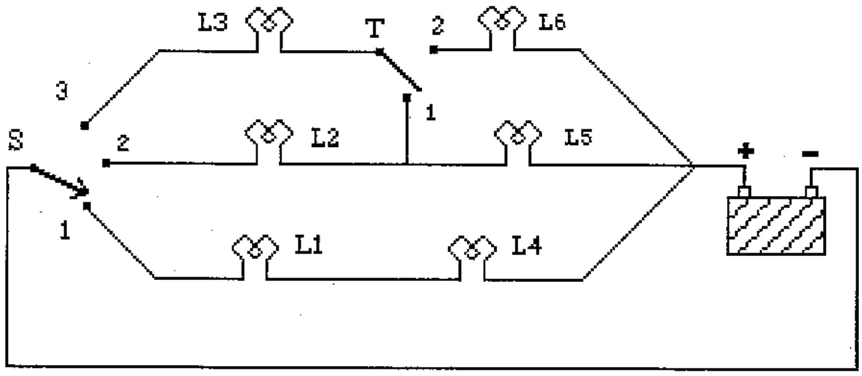
Assertion non calculable à prémisse fausse : circuit de Legrand

XVII. 1999 / XV 2001

Le circuit électrique ci-dessous comporte six lampes identiques notées L1, ..., L6, et deux commutateurs S et T ; S peut prendre trois positions S1, S2, ou S3, et T peut prendre deux positions T1 ou T2.

(a) Que peut-on dire de la vérité de l'assertion suivante
"si L1 est allumée ou si L6 est allumée, alors L3 est allumée ou L4 est allumée" ?

(b) Que peut-on dire de la vérité de l'assertion suivante
"si L1 est allumée et si L3 est allumée, alors L2 est allumée et L5 n'est pas allumée" ?



Question de contrat social (les bonbons de la maîtresse)

XVIII. 1999 / XVI 2001

Dans une école primaire, la maîtresse donne un jour un problème aux élèves, et leur dit : "cherchez ce problème chez vous ; demain, si quelqu'un a su le résoudre, je vous donnerai des bonbons".

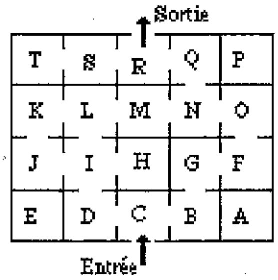
Le lendemain, aucun élève n'a su faire le problème. La maîtresse sort un paquet de bonbons et les distribue aux enfants. Ceux-ci protestent : "Ce n'est pas juste, on n'a pas su faire le problème, on n'a pas droit aux bonbons !". Goguenarde, la maîtresse leur répond qu'elle a parfaitement respecté le contrat ...

Quels commentaires vous inspire ce récit ?

Assertion non calculable à prémisse fausse : labyrinthe de Durand-Guerrier

XIX. 1999 / XVII 2001

Une personne nommée X a traversé le labyrinthe ci-dessous, de l'entrée à la sortie, sans jamais être passée deux fois par la même porte. Les pièces séparées parles portes sont notées A, B,..., T.



Pour chacune des sept phrases suivantes, dire si elle est vraie, si elle est fausse ou si on ne peut pas savoir :

(1) X est passé par T ;
(2) X est passé par N ;
(3) X est passé par M ;
(4) si X est passé par O, alors X est passé par F ;
(5) si X est passé par K, alors X est passé par L ;
(6) si X est passé par L, alors X est passé par K ;
(7) si X est passé par S, il est passé par T.

Assertion calculable à prémisse fausse (fonction)

XVIII. 2001

Un étudiant A affirme que la proposition (I)

(I) : $\forall x \in R$, si $x^2 + 1 \leq 0$, alors$(x^2 + 1)^2 \leq 0$

est vraie, et justifie ainsi son affirmation :

"On a $(x^2 + 1)^2 = x^2(x^2 +1) + (x^2 + 1)$, donc si $x^2 + 1 \leq 0$, alors $x^2(x^2 + 1) \leq 0$, et $(x^2 +1)^2$ est la somme de deux quantités négatives, donc est négative".

Un étudiant B conteste son affirmation ainsi :

"Mais c'est absurde ! L'hypothèse $x^2 + 1 \leq 0$ n'est jamais vraie, ni la conclusion $(x^2 + 1)^2 \leq 0$!De plus, quel que soit le signe de $x^2 + 1$, son carré est évidemment strictement positif ! Cette assertion est donc complètement fausse !"

Que pensez-vous de ce dialogue ... et de la proposition (I) ?

Annexe 3 : les différences femmes (à gauche dans le diagramme)/ **hommes** (à droite)

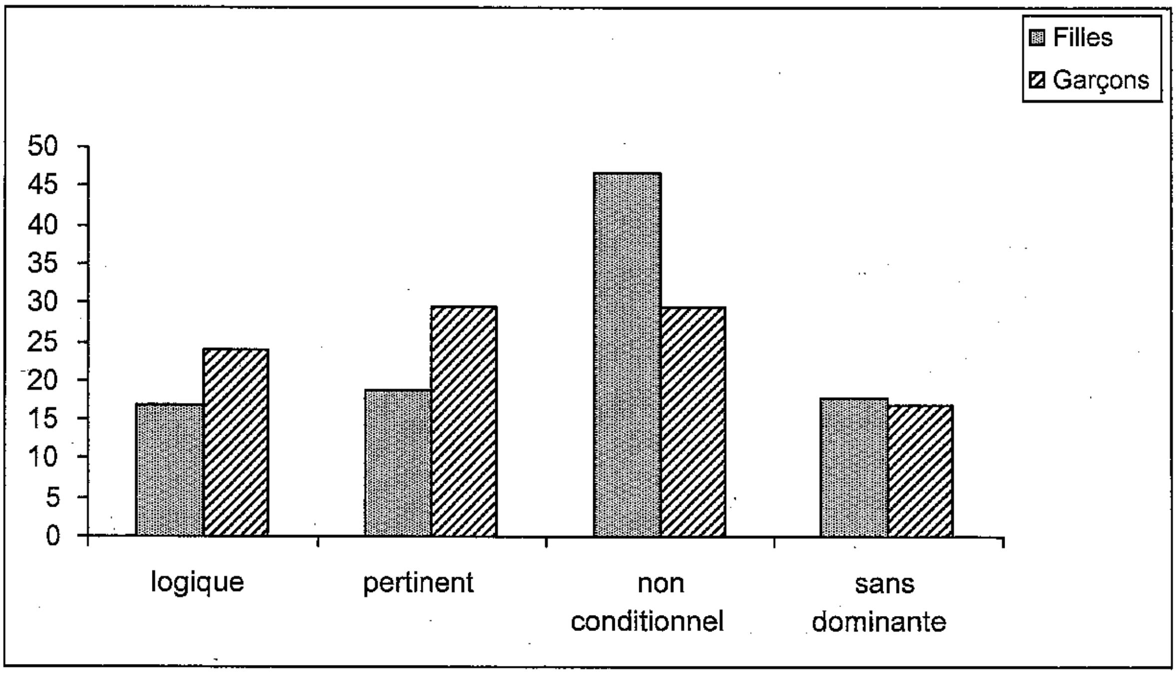


Figure D. *Distribution des profils selon le sexe (test1 + test2 ; N=178)*

Annexe 4. Le faux entraîne le faux, quelques exemples

En psychologie cognitive, les questions sont centrées sur des cas d'inférences où "le vrai entraîne le vrai". Manketow exprime par exemple que "la déduction conserve la vérité" (*Deduction is truth-preserving*), ce qui est exact, mais correspond à une orientation des recherches sur l'implication vers les seules situations où on part du vrai (ou du connu). Mankeltow souligne aussi dans son introduction que "la déduction n'augmente pas l'information: il n'y a pas plus de représentations du monde qui sont exprimées par la conclusion qu'il n'y en avait d'exprimées dans les prémisses. Au mieux, la déduction rend explicite ce qui est déjà là." ("*it does not increase information: no more models are ruled out by the conclusion than were ruled out by the premises. A deduction at best makes explicit what was already there.*". Une telle interprétation exclut les implications à hypothèse fausse et conclusion vraie. (Mankeltow relève par ailleurs le fait qu'on porte aussi des jugements de vraisemblance ; de ce point de vue, il a été montré que l'incertitude sur la vérité des prémisses se transmet aux conclusions).

Ces propriétés avancées pour l'implication peuvent conduire à glisser de "la conservation de la vérité" à"la conservation de la valeur de vérité" : les difficultés observées en programmation dans l'écriture des conditions sous forme de test sur une variable logique, indiquent que même les étudiants de mathématiques ont des problèmes à bien différencier les deux. De là à passer au "théorème en acte" "dans l'implication "le vrai entraîne le vrai et le faux entraîne le faux", il n'y a qu'un pas.

Nous en avons trouvé un bel exemple dans un numéro de"Sciences Humaines" (mai 2001) consacré à l'intelligence. Sur les formes de raisonnement on y trouve (p. 29) une illustration avec des perroquets, et la"légende" suivante :

"Tous les perroquets ont des ailes ; tous les oiseaux ont des ailes ; donc, tous les perroquets sont des oiseaux."Beaucoup de personnes se laissent piéger par ce type de raisonnement qui n'est pas valide logiquement. Par contre, le syllogisme suivant :"Tous les perroquets sont des marins ; tous les marins sont des cachalots ; donc, tous les perroquets sont des cachalots"est un raisonnement parfaitement valide, même si chacune des prémisses qu'il contient est absurde. ***Et donc****, la conclusion aussi. CQFD.*

La lecture du texte lui-même (p. 29) confirme que le "donc" qui relie l'absurdité de la conclusion à celle des prémisses (et donc relève de la conservation de la valeur de vérité : vrai, faux ou absurde !) n'est pas là par hasard :

.. *"Beaucoup de sujets se laissent abuser par le raisonnement suivant : "Tous les perroquets ont des ailes ; tous les oiseaux ont des ailes ; donc, tous les perroquets sont des oiseaux." Car si chacune de ces propositions est empiriquement vraie, le raisonnement, lui, n'est pas valide. Pour se rendre compte de la faille logique, il suffit de remplacer "perroquets" par"mouches". Les mouches ont des ailes, tout comme les oiseaux, mais on ne peut en conclure que les mouches sont des oiseaux ... Inversement, tout le monde admet que le raisonnement suivant conduit à une absurdité :"tous les perroquets sont des militaires ; tous les militaires allaitent leurs petits ; donc, les perroquets allaitent leurs petits."*

Pourtant, le raisonnement lui-même est parfaitement rigoureux. Ce sont les prémisses, et ***partant*** *la conclusion, qui sont fausses*".

Un autre exemple a été relevé dans le traitement par des stagiaires et leur instructeur dans le cadre de la préparation à l'utilisation de la résonance magnétique pour tester la qualité de pièces d'acier. Il s'agit d'une implication physique introduite pour tester l'acquisition des connaissances sur l'aimentation, et plus particulièrement la connaissance que l'existence des pôles d'une barre aimantée est indépendante de la forme de la barre (si elle reste assez "ouverte").

La conditionnelle suivante est introduite : "*si on tord une barre aimantée (sans la chauffer) alors elle garde ses deux pôles*". Il s'agit de choisir la bonne réponse parmi quatre (c'est vrai).

Les stagiaires savent par ailleurs qu'une barre fortement chauffée perd son aimantation.

Les réactions des stagiaires à cette question sont : "*on peut pas*" (tordre une barre sans la chauffer), "*ça n'a pas de sens*", et contestent la pertinence de la réponse dite "correcte". L'instructeur répond en se réfugiant derrière le fait que "tout ça c'est abstrait". Un autre exemple du même ordre est cité par l'auteur qui le traite dans le cadre de sa problématique, qui ne concerne en rien la logique (Delgoudet, C. (2001). La construction des liens entre situations de travail et situations d'apprentissage dans la formation professionnelle. PISTES, 3(2).http://www.unites.uqam.ca/pistes/v3n2/articles/v3n2a2.htm.)