DE L'INFLUENCE DES REPRÉSENTATIONS DISPONIBLES SUR LA RÉSOLUTION DE PROBLÈMES ÉLÉMENTAIRES DE PROBABILITÉ ET SUR L'ACQUISITION DU CONCEPT D'INDÉPENDANCE.

Claire DUPUIS et Suzette ROUSSET-BERT

At the first stages of probability teaching, which changes between which registers of representation can be introduced in order to help understanding ? Modelisations congruent to natural language, the use of so called familiar situations or the restriction to symbolic expressions cannot help pupils to understand, for example, independance in probability. Introducing tree diagrams and cross tables representations in probability is helpful to construct the meaning of probabilistic independance.

Lors de l'initiation aux probabilités, quels sont les changements de registres de représentation à introduire pour faciliter la compréhension des phénomènes et à quel moment les introduire ? Les recours aux modélisations quasi linguistiques, au faux concret évoqué ou aux seules écritures symboliques ne sont pas suffisants pour la compréhension de l'indépendance en probabilité, quand ils ne créent pas eux-mêmes des obstacles. Dès le début, la pratique de changements de registres dans la représentation de situations d'indépendance et de dépendance, utilisant les registres que constituent les arbres et les tableaux de probabilité, permet de donner du sens à l'indépendance en probabilité.

Pour l'avoir observé avec des populations de collégiens et de lycéens, on connaît maintenant l'influence que peut avoir la possibilité de disposer de représentations intermédiaires lorsqu'il s'agit de passer d'un énoncé de problème à sa résolution (par exemple W. L. Damm, C.Dupuis (1992), Groupe Math -Français (1996))

Nous avons réalisé deux questionnaires, le premier en Deug "BioPhysicoChimie" (au début de la deuxième année) et le second en Terminale (après la fin du cours de probabilité). Dans notre analyse, nous distinguerons les questions classiques de calcul de probabilités et celles qui portent sur l'indépendance de deux événements.

Ayant déjà observé les possibilités qu'offrent les arbres de probabilité en terminale (Claire Dupuis et Suzette Rousset-Bert, 1996), nous avons voulu observer si la mise à disposition d'une organisation des données en tableau (ou en arbre) pouvait avoir une influence sur le traitement et sur la réussite de problèmes élémentaires de probabilité. C'est ce que nous étudierons plus particulièrement dans le paragraphe 1, en comparant les réussites des étudiants sur des problèmes équivalents du point de vue de leur contenu mathématique, mais différents dans la présentation des données (énoncés sous forme de texte seulement, énoncés comportant des tableaux ou des arbres). On trouvera en annexe l'ensemble des énoncés ainsi qu'une description des conditions de passation et un aperçu des résultats.

Les paragraphes suivants sont centrés sur l'acquisition du concept d'indépendance et la possibilité d'améliorer l'enseignement de ce concept en faisant travailler les élèves sur ses différentes représentations sémiotiques. Nous ferons tout d'abord un bref panorama des difficultés de présentation de ce concept dans les programmes de Terminale (§2)

© *Annales de Didactique et de Sciences Cognitives*
6 (1998) (p. 67-87) - IREM de Strasbourg

Dans le paragraphe 3, nous montrerons que le concept d'indépendance peut être valablement enseigné et surtout mis en oeuvre dans des registres de représentation autres que la langue naturelle ou les écritures symboliques.

Nous pensons en effet que, pour que les élèves s'approprient rapidement cet objet mathématique, s'en fassent une image mentale correcte, il est bon de rendre "visible" l'indépendance de deux événements dans un arbre et dans un tableau.

Nous analyserons ensuite les réponses des élèves et des étudiants aux questions portant sur l'indépendance.(§4). Pour ce faire, nous définirons diverses conceptions de l'indépendance. Nous nous intéresserons aussi aux représentations sémiotiques, telles que des arbres ou des tableaux, produites spontanément par les élèves et à leur lien avec la réussite aux différentes questions. Nous montrerons que les réussites augmentent lorsque les élèves disposent de plusieurs outils de représentation.

1) LES QUESTIONS "CLASSIQUES" : EFFETS D'UNE VARIATION.

La difficulté du problème ci-dessous, dans sa version "texte", tient au fait que le croisement des deux caractéristiques sémantiques que sont la couleur et la parité n'est pas mis en évidence. Ce croisement serait plus perceptible dans une version "texte" si les trois données étaient des probabilités d'intersections (comme dans le problème T2). Ce croisement est bien mis en évidence par la présentation sous forme de tableau.

Enoncé T1 tableau

Une urne contient des jetons de deux couleurs : JAUNE et VERT, portant chacun un numéro, qui est PAIR ou IMPAIR.

On tire au hasard un jeton dans cette urne.

Le tableau ci-dessous indique les probabilités de différents événements :

	être JAUNE	être VERT	
porter un numéro PAIR	1 / 12		4 / 12
porter un numéro IMPAIR			
	7 / 12		

1) Les événements "être JAUNE" et " porter un numéro PAIR" sont-ils indépendants?
JUSTIFIEZ VOTRE REPONSE.
2) Quelle est la probabilité pour que le jeton soit VERT ?
3) Quelle est la probabilité pour que le jeton porte un numéro IMPAIR ?
4) Quelle est la probabilité pour que le jeton soit VERT et porte un numéro IMPAIR?

Trente-quatre étudiants sur trente-huit ont répondu correctement aux trois questions ne portant pas sur l'indépendance.

Les trente-quatre étudiants qui ont répondu correctement ont tous, quoique ce ne soit pas demandé, complété correctement le tableau. Deux autres étudiants ont complété correctement le tableau, sans répondre explicitement aux questions.

Enoncé T1 texte.

Une urne contient des jetons de deux couleurs : BLANC et ORANGE, portant chacun un numéro.
On tire au hasard un jeton dans cette urne.
La probabilité pour que le jeton soit BLANC est 7 / 12.
La probabilité pour que le jeton porte un numéro PAIR est 4 / 12.
La probabilité pour que le jeton soit BLANC et porte un numéro PAIR est 1 / 12.
1) Quelle est la probabilité pour que le jeton soit ORANGE ?
2) Quelle est la probabilité pour que le jeton porte un numéro IMPAIR?
3) Quelle est la probabilité pour que le jeton soit ORANGE et porte un numéro IMPAIR ?
4) Les événements "être ORANGE" et "porter un numéro IMPAIR" sont-ils indépendants ?
JUSTIFIEZ VOTRE REPONSE.

Vingt-deux étudiants sur trente-huit ont répondu correctement aux trois questions ne portant pas sur l'indépendance.

Représentations spontanées.

Sur les vingt-deux réponses correctes, on observe neuf tableaux complets corrects avec marges, deux arbres corrects et deux arbres où les probabilités des intersections sont sur les branches au lieu d'être au bout. Dix-neuf étudiants ne font aucune représentation, neuf d'entre-eux donnant les trois réponses correctes.

Réussites, échecs et représentations "spontanées".

La différence entre les taux de réussite à T1-Tableau et à T1-Texte est significative et peut être attribuée à la différence de présentation des données des deux versions de T1. En effet, on observera que les deux échantillons de trente-huit étudiants ont exactement les mêmes performances sur les trois autres problèmes.

Dans la population des soixante-seize étudiants interrogés, nous avons observé que

- faire figurer une représentation non demandée sur la copie n'est pas une condition nécessaire de réussite (pour T1-texte, on observe neuf réussites sur vingt-deux sans aucune représentation visible).

- une représentation correcte visible (non demandée) n'est jamais associée à une réponse fausse (elle est associée soit à une réponse correcte, soit à une non-réponse).

- il est rare qu'une représentation non conforme soit associée à une réussite ; ainsi pour T1-texte, les seuls cas sont deux arbres avec les probabilités d'intersection sur les branches au lieu d'être au bout des branches. Compte tenu des questions posées, cette erreur de représentation est sans conséquence sur les réponses.

En conséquence, si la production d'une représentation (non demandée) ne peut être qualifiée de nécessaire à la réussite, on voit que dans tous les cas où cette production spontanée d'une représentation est faite par des élèves qui maîtrisent ce type de représentation, elle conduit à la réussite.

2) L'INDÉPENDANCE EN PROBABILITÉ : QUELQUES AVENTURES DE L'ENSEIGNEMENT DE CET OBJET MATHÉMATIQUE EN TERMINALE

la place de l'indépendance dans l'enseignement des probabilités

Seule l'indépendance de deux événements figure au programme de terminale. Dès que l'on a plus de deux événements, ou que l'on fait des conditionnements successifs, il faut introduire des distinctions entre l'indépendance deux à deux, trois à trois, et l'indépendance mutuelle. On peut penser que le concept d'indépendance figure dans les programmes de Terminale parce qu'il semble simple, du point de vue de l'expert, de l'introduire en même temps que le concept de probabilité conditionnelle. Ce que l'on enseigne à ce niveau n'est donc qu'une toute petite partie de la richesse d'un concept plus utile pour la suite des études, du moins si les probabilités y ont une place, que pour la résolution de problèmes de niveau Terminale. Ainsi, par exemple, l'indépendance des variables aléatoires constituant un échantillon est une des hypothèses permettant de construire la loi de la moyenne d'échantillon $\overline{X}$ qui sera l'outil de décision d'un test d'hypothèses sur une moyenne.

ce qui est fait habituellement dans les manuels pour présenter la notion d'indépendance aux élèves

Nous avons examiné quelques manuels de terminale et leur présentation de l'indépendance ; elle respecte plus ou moins le schéma suivant.

Au paragraphe "indépendance" figure une définition incluant $p(A \cap B) = p(A).p(B)$. Dans certains ouvrages, on démontre l'équivalence avec $p_A(B)=p(B)$ lorsque $p(A)$ n'est pas nul. Il n'y a, en général, pas d'exemple de situation permettant de discerner un cas d'indépendance et un cas de dépendance.

La définition est souvent accompagnée de quelques commentaires du genre "Ainsi la probabilité d'obtenir B sachant que A est réalisé, est égale à la probabilité d'obtenir B. Intuitivement cela signifie que B ne dépend pas de A ; de même A ne dépend pas de B".

Cette tentative louable de rapprochement avec le familier dans la phrase "Intuitivement cela signifie que B ne dépend pas de A" peut conforter des confusions puisqu'elle supprime toute référence à une probabilité dans la définition de l'indépendance. On glisse alors sans le vouloir vers l'idée fausse que l'indépendance est attachée de manière intrinsèque à la définition de l'événement sans tenir compte de la répartition de probabilité sur l'univers.

D'autres manuels essaient d'éclairer cette définition par un exemple supposé familier pour l'élève, celui des tirages successifs de deux boules avec remise.

L'énoncé type de l'exercice proposé pour illustrer l'indépendance est le suivant :

"une urne contient 4 boules blanches et 2 boules noires. On effectue, au hasard, deux tirages successifs (indépendants) avec remise. Quelle est la probabilité pour que la première boule tirée soit blanche et la seconde noire ?"

Si A est l'événement "la première boule tirée est blanche" et si B est l'événement "la seconde est noire", p(AetB) = p(A).p(B) ; $p(A) = \frac{2}{3}$; $p(B) = \frac{2}{6} = \frac{1}{3}$ $\quad p(AetB) = \frac{2}{3} \times \frac{1}{3} = \frac{2}{9}$.

En fait quels sont les événements indépendants dans ce problème et que cache-t-on aux élèves ?

Notons Ω_1 (resp Ω_2) l'ensemble des éventualités associé au premier tirage(respectivement deuxième tirage). A chaque tirage de deux boules avec remise on peut associer un couple de ($\Omega_1 \times \Omega_2$). Si A1 est le sous ensemble de Ω_1 constitué des boules blanches et si A2 est le sous ensemble de Ω_2 constitué des boules noires , alors ce sont les événements $A1 \times \Omega_2$ et $\Omega_1 \times A2$ qui sont indépendants pour la probabilité produit.

Il ne nous semble pas raisonnable de parler aux élèves d'ensembles produit et de probabilité produit ; on parle donc dans cet exemple de l'indépendance de deux événements mal définis. Autrement dit la référence à cet exemple soit disant "familier" est un rideau de fumée.

Regardons maintenant les quelques exercices qui suivent le cours et qui concernent l'indépendance.

Bon nombre d'entre eux contiennent au moins un indice dans l'énoncé qui conduit l'élève inévitablement vers la loi binomiale, comme dans l'énoncé suivant.

" Un représentant de commerce doit visiter 5 clients. Chacune de ces visites est indépendante des autres . La probabilité qu'il rencontre effectivement le client est p=0,6 (le client est absent, ou le représentant ignore le code de l'immeuble).

On note X le nombre de clients effectivement rencontrés. Déterminer la loi de probabilité de X et son espérance mathématique.

Donnez la probabilité des événements suivants :

a) le représentant rencontre au moins un client

b) le représentant rencontre au moins trois clients"

Que doit-on comprendre en lisant la phrase "Chacune de ces visites est indépendante des autres"? L'auteur a voulu exprimer que les cinq variables aléatoires indicatrices X_i associées aux cinq visites de clients sont des variables aléatoires indépendantes de sorte que leur somme X suit une loi binomiale.

On peut imaginer une expérience aléatoire pour donner un sens probabiliste à cette phrase. Le représentant de commerce dispose d'une urne contenant des boules représentant ses différents clients. Il tire 5 boules successivement avec remise. Autrement dit il peut aller voir 5 fois de suite le même client. Il n'y a donc indépendance que pour les représentants de commerce amnésiques.........

Certains auteurs pourraient être tentés de supprimer dans leurs énoncés toute référence à une indépendance de ces variables, arguant du fait que la probabilité de réussir à voir le client est stable. Mais cela ne suffit pas. Il faut en effet se rappeler que la probabilité a priori d'obtenir une boule d'une couleur donnée est aussi stable lorsque l'on tire successivement et sans remise dans une urne à deux couleurs, du moins jusqu'à épuisement de l'urne. Dans ce cas la loi de la somme des X_i est une loi hypergéométrique. Cette suppression de toute référence à l'indépendance ne gênerait guère les élèves qui feraient facilement et éventuellement implicitement cette hypothèse d'indépendance pour construire leur loi binomiale.

L'indépendance intervient aussi dans des exercices qui, à un habillage près, ressemblent à celui-ci : "Monsieur Durand, très prévoyant, a utilisé pour chauffer sa maison deux types de chauffage indépendants : le premier (pompe à chaleur) tombe en panne avec une probabilité de 0,2 et le second (chauffage au fuel) tombe en panne avec une probabilité de 0,05.

Quelle est la probabilité qu'aucun des types de chauffage ne tombe panne? un seul des deux types de chauffage soit en panne ?"

On peut remarquer que ce sont des types de chauffage qui sont qualifiés d'indépendants. On ne sait pas trop ce qu'est l'expérience aléatoire. L'indépendance semble associée de manière intrinsèque aux événements.

On doit comprendre que ce sont les deux événements "le premier chauffage ne tombe pas en panne" et "le second chauffage ne tombe pas en panne" qui sont indépendants pour une probabilité évoquée dans la suite de l'énoncé. L'indépendance de ces deux événements est **postulée a priori,** il s'agit **d'une hypothèse de modélisation.** Il ne reste plus alors qu'à faire le calcul dans le modèle. En fait, même si cette hypothèse ne figurait pas dans l'énoncé les élèves la feraient en acte, d'ailleurs **que pourraient-ils faire d'autre pour résoudre cet exercice** ?

Cet exercice (et tous les autres construits sur le même modèle) donne à la formule $p(A \cap B)= p(A).p(B)$ une valeur de formule générale quels que soient A et B. Cette formule devient une méthode générale, une sorte de théorème en acte qui pourrait s'énoncer ainsi : pour calculer $p(A \cap B)$ il faut faire $p(A).p(B)$.

Le premier contact de l'élève avec l'indépendance est bien souvent réduit à la présentation d'une définition non suivie d'exemple ou suivie d'un exemple qui masque la difficulté. Rien ne permet à l'élève de discerner une situation d'indépendance d'une situation de dépendance. Enfin les exercices qui suivent le cours comportent deux risques :

celui de renforcer une conception intuitive de l'indépendance (indépendance attachée de manière intrinsèque à la définition de l'événement, sans aucune référence à une répartition de probabilité sur un univers)

celui de donner à la formule $p(A \cap B)= p(A).p(B)$ un caractère universel puisque sans elle il serait impossible de traiter l'exercice.

Il n'est donc pas étonnant que les élèves ne gardent qu'une idée un peu floue de l'indépendance, et que ce concept soit mal construit à la sortie de Terminale.

3) CE QUE L'ON PEUT PROPOSER POUR AMÉLIORER LA COMPRÉHENSION DE L'INDÉPENDANCE DE DEUX ÉVÉNEMENTS.

3.1 Comment représenter l'indépendance dans le registre des tableaux ?

Reprenons le problème T1 (énoncé en page 2)

Une urne contient des jetons de deux couleurs : JAUNE et VERT, portant chacun un numéro, qui est PAIR ou IMPAIR. On tire au hasard un jeton dans cette urne.

Voici le tableau complété correspondant à ce problème.

	être JAUNE	être VERT	
porter un numéro PAIR	1 / 12	3 / 12	4 / 12
porter un numéro IMPAIR	6 / 12	2 / 12	8 / 12
	7 / 12	5 / 12	12 / 12

Les événements J "être JAUNE" et P "porter un numéro PAIR" ne sont pas indépendants(question 1).

A partir de ce tableau présentant une situation de non-indépendance, on peut faire construire par les élèves un autre tableau présentant une situation d'indépendance.

	être JAUNE	être VERT	
porter un numéro PAIR	1 / 12	3 / 12	4 / 12
porter un numéro IMPAIR	2 / 12	6 / 12	8 / 12
	3 / 12	9 / 12	12 / 12

On peut visualiser l'indépendance dans ce tableau en repérant que la probabilité située dans chaque case "croisée" est le produit des probabilités situées dans ses marges. La représentation en tableaux illustre donc le fait que s'il y a indépendance des évènements P et J, il y a également indépendance de P et de V (événement contraire de J), de I (événement contraire de P) et J et des événements V et I.

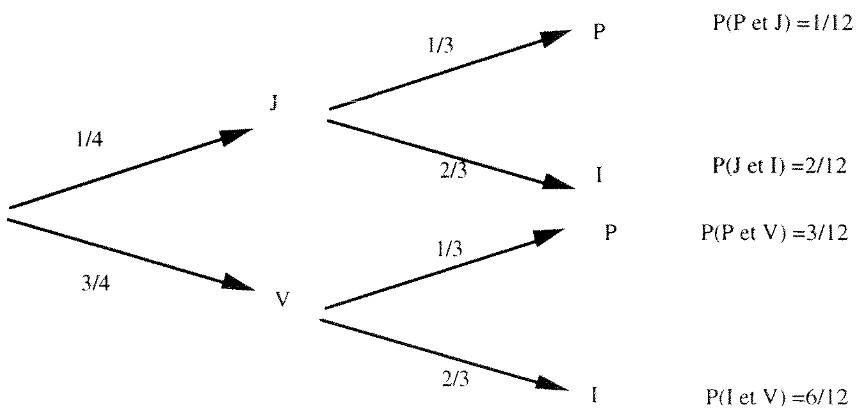
Ce critère de reconnaissance de l'indépendance (le nombre figurant dans chaque case "croisée" est le produit des nombres figurant dans ses marges) est "congruent" à la définition de l'indépendance dans le registre des écritures symboliques ($p(J \cap P) = p(J) \times p(P)$) car il y a correspondance entre les unités signifiantes qui constituent ces deux représentations.

On peut aussi repérer l'indépendance à la proportionnalité des colonnes ou à la proportionnalité des lignes.

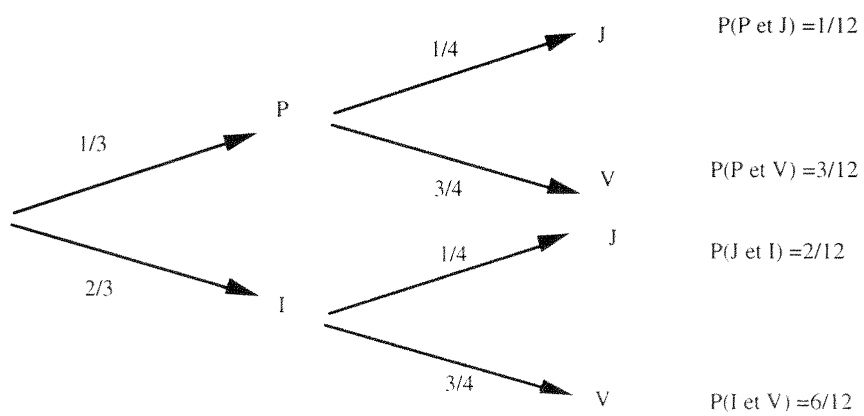
Cette proportionnalité des lignes et colonnes d'un tableau est étudiée en statistiques par les élèves de première ES dans les cas d'hypothèses de proportionnalité parfaite, quand on étudie des problèmes de sous-représentation et sur-représentation.

Comment se traduit l'indépendance dans le registre des arbres ?

Reprenons la situation d'indépendance ci-dessus. Les probabilités des intersections d'événements sont placées aux bouts des branches de l'arbre. Des probabilités conditionnelles apparaissent.



Construisons maintenant l'arbre renversé[1]. Il représente la même situation. Les probabilités aux bouts des branches sont les mêmes.



L'indépendance des événements P et V par exemple se lit de deux façons différentes, toutes deux suffisantes pour établir l'indépendance de deux événements dans cette situation. Sur le

[1] voir Dupuis et Rousset-Bert 1996, Groupe Probabilité 1994.

premier arbre, on voit que les deux "sous-arbres" de deuxième niveau sont identiques. On le voit sur le deuxième arbre aussi, bien évidemment. Si on compare le premier arbre et l'arbre renversé, on voit que le "sous-arbre" de premier niveau se retrouve dupliqué au deuxième niveau de l'autre arbre, et réciproquement !

On voit donc qu'il est facile de représenter une situation d'indépendance de deux événements dans le registre des arbres et dans celui des tableaux.

Avec le même matériel (des boules jaunes et vertes portant des numéros pairs et impairs) on peut donc fabriquer une situation d'indépendance et une situation de dépendance en faisant simplement varier la répartition de probabilité dans l'univers et montrer ainsi que des événements ne sont pas intrinsèquement indépendants ou dépendants.

Le travail dans ces registres de représentations et surtout le travail sur les changements de registres de représentations permettent de donner du sens au concept, par la mise en correspondance des diverses représentations de ce même concept. Les arbres et les tableaux peuvent donc être des aides à la conceptualisation.

4) ÉTUDE DES CONCEPTIONS D'ÉLÈVES DE TERMINALE ET DE DEUG À PROPOS L'INDÉPENDANCE

Nous avons voulu repérer quelles étaient les conceptions des élèves à propos de l'indépendance au travers d'une enquête conçue dans son ensemble pour regarder l'influence de la présentation des données sur les réponses.

Dans ce paragraphe nous avons extrait pour notre analyse les questions concernant l'indépendance dans les différentes problèmes proposés (voir annexes 1 et 2).

4.1 les différentes conceptions des élèves de Terminale

Nous donnons au mot conception le sens que lui a donné Corine Castela (1995) : "Une conception est une modélisation élaborée par le chercheur. Elle vise à rendre compte de la cohérence perceptible dans les comportements observables d'un élève confronté à une groupe de tâches mettant en jeu le même concept mathématique : tout se passe comme si, pour toutes les activités concernées, le sujet utilisait cette conception de l'objet en question. Le modèle proposé n'a pas de visée globale, un même élève peut manifester dans des exercices différents des conceptions distinctes. Il n'est pas non plus strictement local puisqu'il veut traduire une cohérence." Nous définissons les conceptions ci-après.

Conception Correcte : la réponse de l'élève comprend une référence à une des deux définitions $p(A \cap B) = p(A).p(B)$ ou $p_A(B) = p(B)$ et le calcul explicite fait sur l'exemple en question avec sélection des données pertinentes.

Conception fausse incluant des probabilités : la réponse de l'élève comprend une comparaison non pertinente de deux probabilités

ce sont des réponses du type

"les événements jaune et impair sont dépendants car on n'a pas la même probabilité d'avoir un jeton jaune et pair que un jeton vert et impair"

"ils sont dépendants car $p_F(V) \neq p_E(V)$ donc que l'on ait tiré une pièce étrangère ou française on n'a pas les mêmes chances d'avoir une pièce valable" ;

"le fait d'être jaune amène deux possibilités non équiprobables"

"$p(A).p(B) \neq 1$"

Conception intuitive : il n'y a aucune référence à une probabilité dans la réponse.

Dans ces réponses, l'élève fait comme si l'indépendance était une propriété intrinsèque qui qualifie les événements sans tenir compte de la probabilité qui leur est attachée.

Voici quelques exemples :

oui ils sont indépendants car le fait d'être jaune n'a rien à voir avec le fait d'être impair

oui car "valable n'entraîne pas française " et française n'entraîne pas valable"

oui car tous les jetons jaunes ne sont pas pairs

oui car le jeton peut être jaune avec un nombre pair

oui car si on tire un jeton pair, il n'est pas forcément jaune

On peut aussi avec le même type d'arguments obtenir la réponse non :

non car un jeton peut être pair sans être jaune

non car un jeton peut être à la fois jaune et pair.

Des réponses telles que "les événements sont indépendants car ils ne sont pas dépendants" ne nous permettent pas d'imaginer quelle est la conception de l'élève. Il y a ainsi quelques réponses que nous n'avons pas pu interpréter en l'absence d'autres informations.(**Conception non identifiée**).

4.2 les résultats en Terminale

Chacun des quatre problèmes de l'enquête comporte des questions sur l'indépendance. Nous avions conçu ces questions pour regarder l'influence de la présentation des données sur les réponses des élèves.

Dans le problème T1-Tableau, les probabilités à sélectionner pour répondre correctement à la question d'indépendance Q1 sont des données visibles dans le tableau : à la question Q5 portant également sur l'indépendance, il faut se servir de résultats des questions précédentes, également visibles dans le tableau lorsqu'il a été complété correctement. Dans cette enquête, tous les élèves ont complété le tableau correctement.

Dans le problème P'1, la question Q4 porte sur l'indépendance de "la pièce est FRANÇAISE" (F) et "la pièce est VALABLE" (V). La probabilité P(F) est une donnée visible dans l'arbre ; la probabilité P(V) est le résultat de la question précédente. La probabilité de l'intersection P(F et V) est également visible dans l'arbre (en marge).

Dans le problème P'2 , les probabilités à utiliser à la question Q4 sur l'indépendance sont les résultats des trois questions qui précèdent. Il n'y a aucune représentation dans l'énoncé.

Dans le problèmes P''2, les probabilités à utiliser à la question Q6 sur l'indépendance sont des données et des résultats de questions précédentes. Il n'y a aucune représentation dans l'énoncé qui les rende visibles.

	Conception correcte	Comparaison fausse	Conception intuitive	non identifiée	non réponse	nombre d'élèves
T1-Tableau questions Q1 et Q5	**23**	**9**	**23**	**7**	**4**	**66**
P'1 (arbre) question Q4	**23**	**9**	**18**	**7**	**9**	**66**
P'2 (texte) question Q4	**11**	**5**	**5**	**1**	**10**	**32**
P"2 (texte) question Q6	**8**	**2**	**7**	**3**	**14**	**34**

Dans les problèmes T1-Tableau et P'1, les résultats concernant l'indépendance sont stables d'une question à l'autre et il y a très peu de changement de catégorie. Les arguments donnés par les élèves sont de mêmes nature dans les deux problèmes, ils se précisent ou se complètent. Le type de représentation, arbre ou tableau, ne semble pas avoir eu d'influence sur les résultats.

Dans les deux problèmes P'2 et P"2, où aucune représentation n'est fournie, il y a une forte augmentation des non-réponses. On peut penser qu'il y a un problème supplémentaire de sélection des données même pour les élèves qui connaissent la définition correcte et ont su la faire fonctionner dans les problèmes précédents.

4.3 les résultats en Deug

Les étudiants interrogés ont des conceptions stables de l'indépendance et ont répondu de la même façon aux quatre questions sur l'indépendance sauf quatre étudiants pour lesquels nous avons créé la catégorie "Mixte". Dans le tableau ci-après, nous n'avons donc pas distingué les quatre questions mais seulement les deux versions du questionnaire.

Pour interpréter les réponses des étudiants de Deug, nous avons dû introduire des catégories supplémentaires.

Conception correcte avec ambiguïté : la conception est vraisemblablement correcte mais la réponse présente une confusion. Il s'agit soit d'une confusion sur le vocabulaire dépendant, indépendant, soit d'une confusion des symboles $\cup$ et $\cap$ mais l'élève a sélectionné correctement les données pour calculer les probabilités. La catégorie "Correcte avec ambiguïté" diffère peu de la catégorie "Correcte" et il semble qu'il ne s'agisse que de confusion de vocabulaire ou de symbolisme, que l'on peut attribuer au peu de maniement récent de ces notations chez ces étudiants.

Confusion avec l'incompatibilité. Elle se repère dans des réponses telles que : "$p(A \cap B) \neq 0$ donc les événements sont dépendants".

La confusion avec l'incompatibilité ne se rencontre pas en Terminale ce qui n'est guère étonnant puisque la notion d'événements incompatible y est peu travaillée en tant que telle. On rencontre cette réponse un certain nombre de fois en Deug où l'incompatibilité et l'indépendance ont été étudiées l'année précédente.

Conception "Mixte" : l'étudiant utilise au moins une fois une définition correcte et fait au moins une fois une confusion avec l'incompatibilité.

Rappelons qu'il y a 38 étudiants interrogés pour chaque version du questionnaire.

version	Conception correcte	Correcte avec ambiguïté	Confusion avec l'incompatibilité	Conception Mixte	Conception intuitive	non identifiée	non réponse	
A avec T1-tableau	**19**	**5**	**4**	**3**	**3**	**2**	**2**	**38**
B avec T1-texte	**12**	**3**	**3**	**1**	**1**	**4**	**14**	**38**

Les réponses relevant d'une **conception intuitive sont peu nombreuses** ; on les rencontre ici chez des étudiants qui n'ont pas suivi le cursus habituel de la première année et qui n'ont pas suivi de cours de statistique la première année. Le fait d'avoir utilisé l'indépendance en première année de Deug a fait diminuer le nombre de réponses relevant d'une conception intuitive, alors qu'elles étaient assez fréquentes en Terminale

Le nombre de réponses correctes chute au profit des non-réponses entre la version A avec T1-tableau et la version B avec T1-texte, **la question de l'indépendance semblant plus difficile à traiter lorsqu'il n'y a pas de représentation explicite** permettant d'organiser les données et en particulier d'aller repérer la probabilité de l'intersection nécessaire au calcul.

4.4 production spontanée de représentations intermédiaires lorsque les énoncés sont donnés sous forme "texte" ; leur lien avec la réussite aux questions d'indépendance

Nous nous sommes enfin intéressés plus spécifiquement aux représentations non demandées produites spontanément par des élèves et des étudiants dans les problèmes dont l'énoncé est donné en version "texte".

En Terminale, pour résoudre les problèmes P'2 ou P''2, les élèves peuvent évidemment construire une représentation en arbre ou en tableau. Rappelons que les énoncés T1-Tableau et P'1 (arbre) comprennent des représentations. Cinquante-cinq élèves sur soixante-six ont produit spontanément des représentations. Nous avons croisé ces représentations avec les conceptions de l'indépendance.

	Conceptions					
	Correcte	Comparaison fausse	Intuitive	non identifiée	Non réponse	
arbre complet correct	5	3	3		6	17
arbre partiel correct	5		2		2	9
tableau complet correct	4	1		1	1	7
arbre faux	1	2	3	2	12	20
tableau faux				1		1
"patate" fausse					1	1
absence de représentation	4	1	4		2	11
	19	7	12	4	24	66

Sur les dix-neuf réponses relevant d'une conception correcte, on n'observe qu'une représentation fausse. Les vingt et une autres représentations fausses sont associées à des conceptions non correctes ou des non-réponses. **Les représentations correctes spontanées sont majoritairement associées à une conception correcte ou une non-réponse.**

On n'observe que cinq fois une représentation exacte associée à une conception "intuitive", même s'il faut bien admettre qu'une représentation correcte ne suffit pas à remettre en cause une conception intuitive. Seul un travail explicite, tel que celui que nous proposons, d'expression de l'indépendance dans divers registres de représentation pourrait mettre en question cette conception intuitive.

Dans l'enquête **en Deug**, les résultats sont encore plus nets. Considérons les problèmes P1, T1-texte et T2 dont l'énoncé ne comportait pas de représentation. Vingt-deux étudiants ont construit des tableaux corrects non demandés ; ces tableaux sont, pour dix-sept d'entre eux, associés à des réponses relevant d'une conception correcte ou correcte avec ambiguïté, à une confusion avec l'incompatibilité, trois conceptions "Mixtes", une non réponse. Huit étudiants ont construit des arbres corrects non demandés : ils sont associés six fois à des réponses relevant d'une conception correcte ou correcte avec ambiguïté et deux non réponses. **Aucune représentation correcte spontanée n'est associée à une réponse intuitive.**

Il semblerait donc que le fait de produire une représentation correcte spontanée soit associé très majoritairement à un raisonnement probabiliste qui a sa cohérence (même s'il n'est pas exact).

Dans notre enquête de 1995 déjà citée, le problème T1-texte avait été proposé à soixante-huit élèves de Terminale. Seuls trois d'entre eux avaient réussi à donner une réponse à la question de l'indépendance. Nous avions observé beaucoup moins de productions spontanées de représentations : trois arbres exacts, vingt et un arbres diversement faux et un seul tableau. Si on observe plus de représentations non demandées en 1997, c'est peut-être parce que le questionnaire comporte des représentations mais aussi parce que les enseignants en ont plus utilisé dans leurs cours.

Nous pensons qu'il est possible d'améliorer la compréhension de l'indépendance en probabilité en utilisant plus systématiquement les changements de registres de représentation.

CONCLUSION

Tous les éléments recueillis dans nos enquêtes vont dans le même sens : disposer de représentations sémiotiques telles que arbres et tableaux constitue une aide efficace pour la résolution de problèmes et la compréhension des situations probabilistes. Que ces représentations figurent dans l'énoncé ou qu'elles soient produites spontanément par les élèves pour leur propre usage, l'effet est toujours le même : une augmentation des réussites grâce à ces outils d'organisation des données et de segmentation des textes des énoncés.

Les programmes 1998 pour la classe de Terminale ouvrent de nouvelles possibilités. En effet, les arbres et les tableaux sont désormais acceptés comme outils de démonstration, ce qui signifie par exemple pour les arbres que "dans la résolution d'un problème, l'écriture à bon escient d'un arbre pondéré accompagné d'un calcul explicite de la probabilité d'un évènement constitue à elle seule la rédaction d'une justification du résultat obtenu".

Dans le cas particulier de l'indépendance, nous pensons qu'il serait possible d'améliorer encore les performances des élèves en utilisant les repérages visuels de l'indépendance dans les différents registres de représentations sémiotiques. La mise en correspondance des diverses représentations des situations d'indépendance ou de dépendance, dans plusieurs registres différents, permet de donner du sens à un concept difficile à introduire et pour lequel toute tentative de rapprochement avec le familier est plus ou moins vouée à l'échec.

Le concept d'indépendance pourra alors se construire comme un ensemble de représentations sémiotiques intériorisées et c'est la coordination des différents registres de représentations sémiotiques qui va permettre cette construction. Or "La coordination entre des représentations relevant de systèmes sémiotiques différents n'a rien de spontané. Sa mise en place ne résulte pas automatiquement des apprentissages classiques trop directement centrés sur des contenus d'enseignement. Un travail d'apprentissage spécifique centré sur la diversité des systèmes de représentation, sur l'utilisation de leurs possibilités propres, sur leur comparaison par mise en correspondance et sur leurs traductions mutuelles l'un dans l'autre semble nécessaire pour la favoriser. Lorsqu'un tel type de travail est proposé, on constate une modification complète dans les initiatives et dans les démarches des élèves pour effectuer des traitements mathématiques, pour les contrôler , pour la rapidité d'exécution et aussi pour l'intérêt pris à la tâche. Il n'y a pas simplement réussite mais modification de la qualité des productions." (R. Duval, 1995).

RÉFÉRENCES

Corine Castela (1995) Apprendre avec et contre ses connaissances antérieures, *Recherches en Didactique des Mathématiques*, Vol.15, n°1, pp. 7-47.

Werner L. Damm, C.Dupuis (1992) "Les problèmes de mélange" , *Activités Mathématiques*, N° 18, pp 1 - 12.

Claire Dupuis et Suzette Rousset-Bert (1996) Arbres et tableaux de probabilité : analyse en termes de registres de représentation, *Repères-IREM*, N°22, pp. 51-72.

Raymond Duval (1995) *Sémiosis et pensée humaine*, Peter Lang, Bern.

Groupe Math -Français de l'IREM de Strasbourg (1996) *Problèmes de mise en équations : ces charades dont la solution est un système d'équations à deux inconnues*, Brochure IREM de Strasbourg.

Groupe "Probabilité" de l'IREM de Strasbourg (1994) *Enseigner les probabilités en classe de Terminale*, Brochure S 157, IREM de Strasbourg.

ANNEXE 1 LES ÉNONCÉS DU QUESTIONNAIRE DE DEUG

Enoncé P1

En vrac dans un panier se trouvent un grand nombre de pièces de monnaies de différents pays.
Si je prends une pièce au hasard, différents événements peuvent se produire : la pièce peut être française ou étrangère, être encore en cours de validité ou au contraire ne plus être valable.

La probabilité que la pièce soit française et valable est : 41 / 100.
La probabilité que la pièce soit étrangère et valable est : 18 / 100
La probabilité que la pièce soit française et ne soit plus valable est : 32 / 100
La probabilité que la pièce soit étrangère et ne soit plus valable est : 9 / 100

1) Quelle est la probabilité que la pièce soit française ?
2) Quelle est la probabilité que la pièce soit valable ?
3) Les deux événements "la pièce est française" et "la pièce est valable" sont-ils indépendants ? JUSTIFIEZ VOTRE REPONSE.

Enoncé P2

Dans une caisse se trouvent un grand nombre de petites pièces métalliques qui ont été fabriquées dans l'atelier A ou l'atelier B.
Si je prends une pièce au hasard dans la caisse, différents événements peuvent se produire : la pièce peut provenir de l'atelier A ou de l'atelier B, être bien fabriquée ou défectueuse.

Le tableau ci-dessous donne des probabilités d'événements :

	la pièce est bien fabriquée	la pièce est défectueuse
la pièce provient de l'atelier A	1 / 10	2 / 10
la pièce provient de l'atelier B	3 / 10	4 / 10

1) Quelle est la probabilité que la pièce provienne de l'atelier A ?
2) Quelle est la probabilité que la pièce soit bien fabriquée ?
3) Les deux événements "la pièce provient de l'atelier A" et "la pièce est bien fabriquée" sont-ils indépendants ? JUSTIFIEZ VOTRE REPONSE.

Enoncé T2

Une urne contient des jetons de deux couleurs : ROUGE et NOIRE, portant chacun un numéro.
On tire au hasard un jeton dans cette urne.
La probabilité pour que le jeton soit ROUGE et porte un numéro PAIR est 1/9 .
La probabilité pour que le jeton soit ROUGE et porte un numéro IMPAIR est 2/9.
La probabilité pour que le jeton soit NOIR et porte un numéro PAIR est 3/9.
1) Quelle est la probabilité pour que le jeton soit ROUGE ? soit NOIR ?
2) Quelle est la probabilité pour que le jeton porte un numéro PAIR ? IMPAIR ?
3) Quelle est la probabilité pour que le jeton soit NOIR et porte un numéro IMPAIR ?
4) Les événements "être NOIR" et " porter un numéro IMPAIR" sont-ils indépendants ?
JUSTIFIEZ VOTRE REPONSE.

ANNEXE 2 LES ÉNONCÉS DU QUESTIONNAIRE DE TERMINALE

Le questionnaire de Terminale comprenait deux modalités :

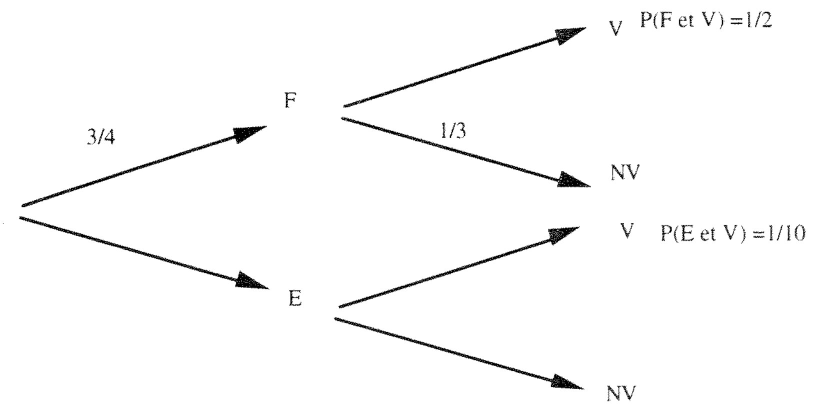
modalité A	T1 tableau	P'1	P'2
modalité B	T1 tableau	P'1	P"2

Enoncé T1 tableau

Même début d'énoncé qu'en DEUG avec deux questions supplémentaires :
5) Les événements "être VERT" et " porter un numéro IMPAIR" sont-ils indépendants ?
JUSTIFIEZ VOTRE REPONSE.
6) Quelle est la probabilité pour que le jeton soit JAUNE sachant qu'il porte un numéro PAIR ?

Enoncé P'1

En vrac dans un panier se trouvent un grand nombre de pièces de monnaies de différents pays.
Je prends une pièce au hasard dans le panier.
Différents événements peuvent se produire : la pièce peut être FRANCAISE (F) ou ETRANGERE (E), être encore VALABLE (V) ou au contraire ne PLUS être VALABLE (NV).
L'arbre ci-dessous donne les probabilités de différents événements :



1) Quelle est la probabilité que la pièce soit VALABLE sachant qu'elle est FRANCAISE ?
2) Quelle est la probabilité que la pièce soit FRANCAISE et ne soit PLUS VALABLE ?
3) Quelle est la probabilité que la pièce soit VALABLE ?
4) Les deux événements "la pièce est FRANCAISE" et "la pièce est VALABLE" sont-ils indépendants ?
JUSTIFIEZ VOTRE REPONSE.
5) Quelle est la probabilité que la pièce soit FRANCAISE sachant qu'elle est VALABLE ?

Enoncé P'2

Dans une caisse se trouvent un grand nombre de petites pièces métalliques qui ont été fabriquées dans l'ATELIER A ou l'ATELIER B.
Je prends une pièce au hasard dans la caisse.
Différents événements peuvent donc se produire : la pièce peut provenir de l'ATELIER A ou de l'ATELIER B, être BIEN FABRIQUEE ou DEFECTUEUSE.

La probabilité pour que la pièce provienne de l'ATELIER B est 15 / 25.
La probabilité pour que la pièce soit DEFECTUEUSE est 3 / 25.

La probabilité pour qu'à la fois la pièce provienne de l'ATELIER B et soit DEFECTUEUSE est 2 / 25.

1) Quelle est la probabilité que la pièce provienne de l'ATELIER A ?
2) Quelle est la probabilité que la pièce soit BIEN FABRIQUEE ?
3) Quelle est la probabilité que la pièce provienne de l'ATELIER A et soit BIEN FABRIQUEE ?
4) Les deux événements "la pièce provient de l'ATELIER A" et "la pièce est BIEN FABRIQUEE" sont-ils indépendants ? JUSTIFIEZ VOTRE REPONSE.
5) Quelle est la probabilité pour que la pièce provienne de l'ATELIER B sachant qu'elle est DEFECTUEUSE ?

Enoncé P''2

Dans une caisse se trouvent un grand nombre de petites pièces métalliques qui ont été fabriquées dans l'ATELIER A ou l'ATELIER B.

Je prends une pièce au hasard dans la caisse.
Différents événements peuvent donc se produire : la pièce peut provenir de l'ATELIER A ou de l'ATELIER B, être BIEN FABRIQUEE ou DEFECTUEUSE.

La probabilité pour que la pièce provienne de l'ATELIER B est 15 / 25.
La probabilité pour qu'à la fois la pièce provienne de l'ATELIER B et soit DEFECTUEUSE est 2 / 25.
La probabilité pour que la pièce soit DEFECTUEUSE sachant qu'elle provient de l'ATELIER A est 1 / 10.

1) Quelle est la probabilité que la pièce provienne de l'ATELIER A ?
2) Quelle est la probabilité pour que la pièce soit DEFECTUEUSE sachant qu'elle provient de l'ATELIER B ?
3) Quelle est la probabilité pour qu'à la fois la pièce provienne de l'ATELIER A et soit DEFECTUEUSE ?
4) Quelle est la probabilité pour que la pièce soit DEFECTUEUSE ?
5) Quelle est la probabilité pour que la pièce provienne de l'ATELIER B sachant qu'elle est DEFECTUEUSE ?
6) Les deux événements "la pièce provient de l'ATELIER B" et "la pièce est DEFECTUEUSE" sont-ils indépendants ? JUSTIFIEZ VOTRE REPONSE.

ANNEXE 3 PASSATION ET APERÇU DES RÉSULTATS

Tous les problèmes des questionnaires utilisent des situations d'un tirage au hasard dans une "urne" où les "boules" présentent deux caractéristiques croisées. Les énoncés ont été construits en faisant varier la présentation des données de la situation pour observer si cette variation avait une influence sur les réponses.

Les questionnaires comportent chacun deux modalités : des problèmes communs aux deux modalités permettent de contrôler l'homogénéïté des deux sous-échantillons.

1) Le questionnaire en DEUG.

Nous avons repris les situations des énoncés T1 (T1 texte ici) et T2 de notre article (C. Dupuis et S. Rousset-Bert, 1996). Nous avons modifié les couleurs des boules et les valeurs numériques dans T1 puisque les deux énoncés T1 et T2 sont présents dans les deux versions du questionnaire en DEUG. Nous n'avons conservé que les questions portant sur des probabilité d'intersections, des probabilités marginales et sur l'indépendance car les probabilités conditionnelles ne sont pas utilisées dans le cours de statistique de ces étudiants.

En terminale, à la fin du cours de probabilité (C. Dupuis et S. Rousset-Bert, 1996), nous avions observé

T1 Questions 1 et 2 (probabilités marginales)	65 réponses correctes pour 68 élèves
T1 Questions 3 (probabilité d'intersection)	6 réponses correctes pour 68 élèves !!! (et 31 réponses où la probabilité de l'intersection était calculée par produit des deux probabilités marginales précédentes).
T2 Questions 1 et 2 (probabilités marginales)	52 réponses correctes pour 63 élèves
T2 Questions 3 (probabilité d'intersection)	33 réponses correctes pour 63 élèves.

Deux versions de questionnaire ont été réalisées en DEUG.

	première page	deuxième page	troisième page	quatrième page
version A	T1 tableau **tableau incomplet avec marges visualisées**	P2 **tableau complet sans marges visualisées**	T2 **texte**	P1 **texte**
version B	P1 **texte**	T2 **texte**	T1 texte **texte**	P2 **tableau complet sans marges visualisées**

Conditions de passation.

Les deux versions de questionnaire ont été distribuées dans un amphi de BioPhysicoChimie 2ème année au début de l'année universitaire (durée de passation : une heure de cours). Ces étudiants (sauf ceux qui sont arrivés par équivalence mais qui ne représentent que 11 sur 76) ont eu en première année un cours semestriel de probabilité-statistique (42 heures en tout) ; l'enseignement des probabilités y est conçu comme devant apporter les outils nécessaires à l'apprentissage de l'estimation par intervalle et des tests d'hypothèses qui sont le véritable but du cours. Quasiment tous les énoncés de problèmes y sont des énoncés de modélisation, analogues à ceux que nous avons présenté dans le questionnaire. L'enseignement de deuxième année qui vient de débuter au moment de la passation est un enseignement de tests statistiques. Les tableaux qui y sont utilisés sont des tableaux de données d'observation. Trente-huit étudiants ont eu la version A, et trente-huit autres la version B.

Représentations fournies et réussites aux questions "classiques" avant la question sur l'indépendance

Problème	données et questions	Version A : 38 étudiants	Version B : 38 étudiants
Problème P1	4 probabilités d'intersection 2 questions	**37** **texte**	**37**
Problème P2	4 probabilités d'intersection 2 questions	**34** **tableau complet sans marges**	**34**
Problème T2	3 probabilités d'intersection 3 questions	**33** **texte**	**33**

Réussites, échecs et représentations visibles.

Problème P2.

Le problème est exactement le même dans les deux versions. La seule différence est la place dans le questionnaire. On n'observe pas de différence significative entre les deux échantillons d'étudiants en ce qui concerne les représentations visibles. La plupart (65 sur 76) des étudiants ne font pas de représentation pour répondre à cette question facile pour eux. Cette observation permet de vérifier que sur ces notions de base les deux échantillons d'étudiants sont comparables.

Problème P1.

Le problème est exactement le même dans les deux versions. La seule différence est la place dans le questionnaire : dans la version A, ce problème vient en dernier, après deux tableaux et un énoncé "texte" ; dans la version B, cet énoncé est le premier. La différence entre les deux versions se situe au niveau des représentations visibles : il y a beaucoup plus de tableaux dans la version A que dans la version B ! Mais il n'y a pas, pour autant, de différences de taux de réussites pour ces questions très faciles.

L'étudiant de la version A qui échoue à ce problème a donné des réponses fausses ou n'a pas répondu aux problèmes P2 et T2. Pour T1 tableau, il a complété correctement le tableau sans reprise des réponses en dehors du tableau.

L'étudiant de la version B qui échoue à ce problème calcule souvent les probabilités marginales par des produits mais parfois par des sommes ou des différences. Il a donc quelques réponses correctes mais ne réussit complètement aucun problème.

Croisement des réponses à P1 et P2

Les 68 étudiants qui réussissent P2 réussissent aussi P1. La réussite à P2 impliquant donc la réussite à P1, on peut dire que le problème P2 est plus difficile que P1.

Voisinage

Il n'y a pas d'influence de "voisinage" sur les réussites puisque les problèmes P1, P2 et T2 sont réussis de la même façon quelle que soit leur place dans le questionnaire.

Croisement des réponses à T1-texte et T2 (version B)

Les vingt-deux étudiants qui ont réussi T1-texte ont tous réussi T2. La réussite à T1-texte impliquant donc la réussite à T2, on peut dire que le problème T1-texte est plus difficile que T2.

T1 et T2

Comme en terminale lors de l'enquête précédente, T1-Texte reste un problème plus difficile que T2 sauf lorsque la présentation des données sous forme de tableau dans T1-Tableau met en évidence le croisement des deux caractéristiques (ou déterminations sémantiques) des individus. On voit ici la puissance de cet outil d'organisation des données lorsqu'il y a une difficulté.

2) Le questionnaire de Terminale

En Terminale, les questions étaient plus nombreuses et plus variées et les erreurs sont plus diverses. Nous n'analyserons pas ici ces réponses en détail, nous limitant à une question difficile dans chaque problème, "difficile" parce que sa réponse nécessite d'utiliser des quantités qui ne sont pas toutes des données.

forme du problème	données et questions	
Problème T1 - Tableau tableau incomplet avec marges	2 probabilités marginales et 1 probabilité d'intersection question Q4	**54 réussites sur 66 élèves (0,82)**
Problème P'1 arbre	1 probabilité marginale 2 probabilités d'intersection 1 probabilité conditionnelle question Q5	**39 réussites sur 66 élèves (0,59)**
Problème P'2 texte	2 probabilités marginales 1 probabilité d'intersection question Q3	**15 réussites sur 32 élèves (0,47)**
Problème P''2 texte	1 probabilité marginale 1 probabilité d'intersection 1 probabilité conditionnelle question Q5	**15 réussites sur 34 élèves (0,44)**

Le phénomène persistant observable est, qu'**en cas de difficulté, on retrouve une plus grande réussite lorsque l'énoncé comporte un outil d'organisation.**