Université de Provence

Département des sciences de l'éducation

Année 2007-2008 – Master 1

UE SCER22 :

« Didactique des disciplines scientifiques »

Yves Chevallard
y.chevallard@aix-mrs.iufm.fr

Trois leçons sur la didactique des PER

➔ Leçon 3

Sommaire. – 0. De $R^{\lozenge}$ à $R^{\heartsuit}$ au service d'un projet : deux exemples / 1. Excrire : suite / 2. Excrire : généralisation / 3. Se former à l'enquête codisciplinaire / Annexe

0. De $R^{\lozenge}$ à $R^{\heartsuit}$ au service d'un projet : deux exemples

a) Lancer et conduire d'un PER, c'est piloter un système didactique du type $S(X ; Y ; Q)$, dont le bilan du fonctionnement pourra s'écrire ainsi :

$$[S(X ; y ; Q) \Rightarrow \{ R_1^{\lozenge}, R_2^{\lozenge}, \ldots, R_n^{\lozenge}, O_{n+1}, \ldots, O_m \}] \hookrightarrow R^{\heartsuit}.$$

On a résumé un tel parcours en quelques « gestes » de base, qui consistent à ***observer***, ***analyser***, ***évaluer*** des réponses $R^{\lozenge}$ et à ***développer*** puis ***défendre & illustrer*** sa réponse $R^{\heartsuit}$.

b) Revenons à la question notée Q_0 lors de la leçon 1.

> Q_0. Que sont les praxéologies dont les auteurs cités dans la leçon 1 à propos de pourcentages et autres proportions se prévalent dans leurs observations critiques à l'encontre des écrits ou propos dont ils s'y font l'écho ?

➔ On a évoqué, dans la leçon 2, plusieurs réponses $R^{\lozenge}$ possibles, en les analysant et en les évaluant sommairement : la technique « par le multiplicateur » a ainsi été évaluée plutôt négativement ***dans le contexte considéré***.

➔ Imaginons ici que la question Q_0 ait émergé dans le cadre d'un ***projet*** d'un professeur de mathématiques désireux de présenter à des élèves de 2[de] des textes du site Pénombre sur les pourcentages, accompagnés de commentaires mathématiques, avec l'idée de demander à ses élèves de ***continuer*** sur d'autres textes présents sur ce même site le travail amorcé par lui.

➔ On peut imaginer que, après différentes explorations, le professeur ait esquissé une réponse $R^{\heartsuit}$ qui le conduise à proposer pour le texte ci-après les commentaires qui le suivent.

Texte 0

Quintuplés

Dans la Croix du mardi 6 juillet 1999, Bruno Chenu, dans un article intitulé « Les millions du foot » écrit : « Ce qui se passe depuis un mois dans la rubrique des transferts donne le vertige. En quatre ans en France, le montant des transactions a quadruplé : 400 % d'inflation, ce n'est pas mal ! »

Le vertige, et… la berlue, pour prendre plus 400 %, qui est un quintuplement, pour seulement un quadruplement. En effet, plus 100 % est un doublement, donc plus 200 % un triplement…

Commentaire

Lorsqu'une grandeur variable G (ici, le « montant des transactions ») augmente de r % en passant d'une valeur g_0 à une valeur g_1, on a par définition :

$$g_1 = g_0 + \frac{r}{100}\, g_0.$$

Si $g_1 = 4\, g_0$, on a

$$4\, g_0 = g_0 + \frac{r}{100}\, g_0.$$

En simplifiant par g_0 (supposé non nul !), on obtient

$$4 = 1 + \frac{r}{100}$$

ce qui donne $r = 300$. Lorsque une grandeur quadruple, elle augmente de 300 %, non de 400 %. Inversement, si une grandeur augmente de 400 %, on a

$$g_1 = g_0 + \frac{400}{100}\, g_0 = g_0 + 4\, g_0 = 5\, g_0.$$

La grandeur quintuple.

Au niveau de la vie

Plus récemment, dans Télérama du 29 décembre 1999, Gérard Chaliand, expert en géostratégie, à la suite d'une question sur les résultats des élections législatives de Russie de décembre 1999, interprète celles-ci comme « le symptôme d'une population désorientée », et ajoute : « Comment réagirions-nous si nous avions perdu entre 300 et 400 % de notre niveau de vie en moins de cinq ans ? »

On sait qu'en Russie les températures descendent souvent en dessous de zéro, on ne savait pas qu'il pouvait en être de même du niveau de vie !

Commentaire

Lorsqu'une grandeur G (ici, le « niveau de vie ») diminue de r % en passant d'une valeur g_0 à une valeur g_1, on a par définition :

$$g_1 = g_0 - \frac{r}{100}\, g_0.$$

Lorsque, par définition, la grandeur G ne peut avoir que des valeurs positives ou nulles, soit lorsqu'on a toujours

$$g_1 \geq 0$$

on a

$$g_0 - \frac{r}{100}\, g_0 \geq 0.$$

En divisant par g_0 (supposé non nul), on obtient

$$1 - \frac{r}{100} \geq 0$$

soit $\frac{r}{100} \leq 1$, ou $r \leq 100$. En conséquence, G ne peut diminuer de plus de 100 % : on ne peut parler d'un niveau de vie diminuant de 300 %, par exemple. Si G diminuait de 300 % à partir de la valeur g_0, on aurait

$$g_1 = g_0 - \frac{300}{100} g_0 = g_0 - 3\, g_0 = -\, 2\, g_0$$

ce qui n'a pas de sens en la matière…

➔ On pourra de même imaginer que, alors, le professeur demande à ses élèves, à titre de ***travail hors classe***, d'élaborer sur ce patron des commentaires mathématiques appropriés pour les trois textes suivants : nous n'irons pas plus loin sur cette voie.

Texte 1

Un pourcentage de 100 % ou plus est difficilement compréhensible. Il vaudrait mieux parler, à partir de là, de doublement, de triplement etc. De même, quand la baisse est importante, de parler d'une division par 2, plutôt que d'une baisse de 50 %, d'une division par 3, plutôt que d'une baisse de 66,6666666… %.

Commentaire

Texte 2

Une croissance des prix, qui s'exprime normalement en pourcentage, les prix étant supposés évoluer lentement, peut toujours devenir inflationniste, par exemple atteindre 8 000 %. Ce qui fait une multiplication par combien, cher lecteur?

Commentaire

Texte 3

Dans un article du Figaro au sujet d'un nouveau revêtement routier anti-dérapant : « Le résultat est encore plus significatif à 50 km/h : il faut à peine 10 mètres pour s'arrêter en freinage d'urgence sur ce substrat ultra-râpeux, contre plus de 14 mètres sur une chaussée classique, ce qui représente une diminution de plus de 40 % de la distance d'arrêt. »

Commentaire

➔ Dans ce qui précède, on a vu (ou aperçu) certaines praxéologies ***inscrites*** dans des textes – dans ceux du site Pénombre, dans les « œuvres » $O^{\%}_{\mathrm{Wf}}$ et $O^{\%}_{\mathrm{We}}$. On s'est efforcé aussi de les « dé-inscrire » ou, comme on l'a dit, de les ***excrire***. Enfin, la praxéologie résultant de ce travail a été ***réinscrite*** dans le commentaire donné plus haut. Ainsi a-t-on vu fonctionner, en modèle réduit sans doute, la ***dialectique de l'excription et de l'inscription***.

c) Un parcours d'étude et de recherche, un PER, n'est pas toujours aussi « rapide ». On a lancé dans la leçon 2 un autre PER, impulsé par la question $Q^{\#}$ suivante : « Comment analyser des données ? » Cette question, sur laquelle nous allons revenir, a conduit à étudier un texte – un chapitre d'un ouvrage récent de Philippe Cibois – dont l'examen a suscité une question auxiliaire Q_{χ^2} : « Qu'est-ce que le test du χ^2 ? » À cette question, nous avons imaginé que, après avoir enquêté dans le cadre d'un PER intéressant d'abord le seul système didactique qu'on peut noter $S(y ; y ; Q_{\chi^2})$, y se présente devant X avec un ***compte rendu d'enquête*** – un « cours » – où il aura inscrit une certaine réponse $R^{\heartsuit}$ (voir la leçon 2, 1, e).

➔ Très souvent, en fait, l'enquête réalisée, qui fournit une ***première*** réponse, doit être ***étendue***, ou ***approfondie***, car, en même temps qu'elle répond à des questions, elle en soulève d'autres. Ici, par exemple, *y* peut évaluer comme insuffisamment clarificateurs les éléments de son compte rendu d'enquête relatifs au passage suivant extrait de l'ouvrage *Les méthodes d'analyse d'enquêtes* : « On se souviendra que le khi-deux étant sensible aux effectifs, dès qu'une population d'enquêtés devient importante, il devient rare que le khi-deux d'un tableau croisé ne soit pas significatif. » Que signifie l'assertion selon laquelle le test du χ^2 est « sensible aux effectifs » ? Il semble que l'auteur ait inscrit dans cette affirmation des éléments praxéologiques que, jusqu'ici, nous n'avons pas excrits. On peut donc souhaiter sur ce point un complément d'enquête – une ***reprise de l'enquête***.

➔ Dans une présentation en ligne du n° 69 du *Bulletin de méthodologie sociologique* paru en janvier 2001 (http://www.cmh.pro.ens.fr/bms/abstracts/N.69.htm), on trouve le résumé d'un article de Philippe Cibois intitulé « Faire comprendre le Khideux. À propos de l'*Introduction aux méthodes statistiques en sociologie* de Thierry Blöss et Michel Grossetti ». Le voici.

> À propos d'un ouvrage récent d'introduction aux statistiques, on se pose des questions sur l'enseignement du test du Khideux. Il est proposé de ne pas appliquer le calcul du test comme une simple recette mais d'aider les étudiants à comprendre que le Khideux est homogène à un effectif et que les précautions à prendre pour son emploi relèvent simplement de la prudence.

On se trouve à nouveau devant une ***zone grise***, celle dessinée par l'assertion « le Khideux est homogène à un effectif », sans parler du lien entre cette propriété supposée de l'indicateur qu'est le χ^2 et le fait que « les précautions à prendre pour son emploi relèvent simplement de la prudence ».

➔ On peut chercher d'autres textes où le même auteur développerait un tant soit peu sa pensée. On trouve ainsi en ligne un texte intitulé *Les écarts à l'indépendance. Techniques simples pour analyser des données d'enquête* (http://www.scienceshumaines.com/textesInedits/Cibois.pdf) sur lequel on s'arrêtera un peu plus loin, après s'être remis en mémoire l'expression du χ^2 vue lors de la leçon 2 :

$$\chi^2 = \sum_{i,j} \frac{(n_{ij} - c_{ij})^2}{c_{ij}}.$$

Rappelons que n_{ij} est ici l'effectif ***observé*** de la cellule (i, j) située au croisement de la i-ième ligne et de la j-ième colonne, tandis que c_{ij} est l'effectif ***théorique*** de la même cellule. La contribution au χ^2 de la cellule (i, j) est égal à

$$\frac{(n_{ij} - c_{ij})^2}{c_{ij}}.$$

Si l'on note $e_{ij} = n_{ij} - c_{ij}$ ce qui sera appelé ci-après l'écart « brut, ordinaire » de la cellule (i, j), on voit que « le χ^2 » de la cellule (i, j) peut s'écrire

$$\frac{e_{ij}^2}{c_{ij}} = e_{ij} \times \frac{e_{ij}}{c_{ij}}$$

Lisons maintenant le passage annoncé en incluant entre crochets des « commentaires » mathématiques appropriés.

Cet écart pondéré [à savoir $e_{ij} \times \frac{e_{ij}}{c_{ij}}$, où e_{ij} est l'écart et où $\frac{e_{ij}}{c_{ij}}$ est le poids ou coefficient de pondération] est appelé le khi-deux d'une case. Le créateur de cette indice, Karl Pearson a employé pour le désigner une lettre grecque mise au carré, χ^2, car cet indice est souvent présenté sous la forme suivante où un carré intervient :

$$(\text{observé} - \text{théorique})^2/\text{théorique}$$

$$[\text{soit } \frac{(n_{ij} - c_{ij})^2}{c_{ij}}]$$

or « observé – théorique » [soit $n_{ij} - c_{ij}$] correspond à l'écart à l'indépendance [soit e_{ij}] et le khi-deux peut s'écrire :

$$\text{écart}^2 / \text{théorique}$$

$$[\text{soit } \frac{e_{ij}^2}{c_{ij}}]$$

ou encore :

$$\text{écart} \times (\text{écart} / \text{théorique})$$

$$[\text{soit } e_{ij} \times \frac{e_{ij}}{c_{ij}}]$$

c'est-à-dire l'écart brut à l'indépendance [soit e_{ij}] pondéré par le rapport écart / théorique [soit $\frac{e_{ij}}{c_{ij}}$].

Le khi-deux d'une case [soit $\frac{e_{ij}^2}{c_{ij}}$] est la même chose que l'écart pondéré [soit $e_{ij} \times \frac{e_{ij}}{c_{ij}}$] : on peut ainsi considérer le khi-deux de toutes les cases d'un tableau comme la somme des écarts pondérés [soit la somme des $e_{ij} \times \frac{e_{ij}}{c_{ij}}$, pour tout i et tout j], la somme des « informations de surprise » faibles ou fortes apportées par les différents écarts du tableau. Un écart [soit e_{ij}] apporte toujours une certaine information, mais pour l'indice khi-deux, elle est pondérée par son rapport au théorique [soit par $\frac{e_{ij}}{c_{ij}}$] qui l'amplifie ou la réduit selon le cas. On passe de l'écart brut, ordinaire [soit e_{ij}], à l'information apportée par cet écart [soit $e_{ij} \times \frac{e_{ij}}{c_{ij}}$], notée en nombre pondéré [par $\frac{e_{ij}}{c_{ij}}$] d'individus en écart à l'indépendance [soit e_{ij}], avec donc une unité de compte homogène à un écart [à savoir e_{ij}] (et comparable à lui [e_{ij}] de ce fait).

➔ Ce qui précède peut sembler obscur, parce que l'auteur s'est refusé à employer les notations mathématiques utiles. Pour clarifier les choses davantage, on peut d'abord noter que l'écart $e_{ij} = n_{ij} - c_{ij}$ est positif, négatif ou nul, ce que tout ce qui précède semble ignorer. En réalité, si l'on pose $\varepsilon_{ij} = |e_{ij}| = |n_{ij} - c_{ij}|$, on voit que l'on a aussi

$$\frac{e_{ij}^2}{c_{ij}} = \frac{(n_{ij} - c_{ij})^2}{c_{ij}} = \frac{|n_{ij} - c_{ij}|^2}{c_{ij}} = \frac{\varepsilon_{ij}^2}{c_{ij}} = \varepsilon_{ij} \times \frac{\varepsilon_{ij}}{c_{ij}}.$$

La contribution de la cellule (i, j) au χ^2 du tableau est donc égale à l'écart « absolu » $\varepsilon_{ij} = |n_{ij} - c_{ij}|$ multiplié par le coefficient de pondération $\frac{\varepsilon_{ij}}{c_{ij}}$. Comment alors traduire la notion d'homogénéité évoquée ? C'est ce que l'on va voir maintenant.

d) Imaginons un tableau T de dimensions quelconques, dont l'effectif figurant dans la cellule (i, j) est noté n_{ij} ; on suivra l'argument présenté ci-après sur le cas particulier d'un tableau très simple, à deux lignes et trois colonnes, que voici :

	A	B	C	Totaux
1	**32**	**11**	**20**	
2	**27**	**8**	**25**	
Totaux				

➔ Connaissant les effectif n_{ij} on peut calculer les effectifs marginaux (et l'effectif total) en sommant les colonnes pour obtenir les effectif marginaux $n_{\bullet j}$ par les formules

$$n_{\bullet j} = \sum_i n_{ij}$$

et les effectifs marginaux $n_{i\bullet}$ par les formules $n_{i\bullet} = \sum_j n_{ij}$. On a de plus pour l'effectif total de l'échantillon : $N = \sum_i n_{i\bullet} = \sum_j n_{\bullet j}$. Dans le cas particulier suivi, il vient ceci.

	A	B	C	Totaux
1	32	11	20	**63**
2	27	8	25	**60**
Totaux	**59**	**19**	**45**	**123**

➔ Cela fait, on peut calculer les effectifs théoriques par la formule

$$c_{ij} = \frac{n_{\bullet j} \times n_{i\bullet}}{N}.$$

On a en l'espèce ceci.

	A	B	C	Totaux
1	**30,22**	**9,73**	**23,05**	63
2	**28,78**	**9,27**	**21,95**	60
Totaux	59	19	45	123

On peut alors calculer les écarts $e_{ij} = n_{ij} - c_{ij}$:

	A	B	C
1	**1,78**	**1,27**	**–3,05**
2	**–1,78**	**–1,27**	**3,05**

➔ Avant d'aller plus loin, demandons à un logiciel en ligne que nous avons déjà utilisé (http://www.physics.csbsju.edu/stats/contingency_NROW_NCOLUMN_form.html) de nous dire ce qu'il en est de la valeur de χ^2 et de la probabilité d'obtenir, sous l'hypothèse H_0 d'indépendance des deux caractères croisés, une valeur aussi élevée que cela. On obtient ceci.

$r \times c$ Contingency Table: Results

The results of a contingency table X^2 statistical test performed at 14:09 on 31-MAR-2008

```
data: contingency table

       A       B       C

1      32      11      20      63
2      27       8      25      60

       59      19      45     123



expected: contingency table

        A          B          C

1    30.2       9.73       23.0
2    28.8       9.27       22.0
```

chi-square = 1.38
degrees of freedom = 2
probability = 0.501

La valeur de χ^2 est d'environ 1,38 ; la probabilité de l'événement $\chi^2 \geq 1{,}38$ si H_0 est vraie est voisine de 0,5 : l'événement observé ***n'a donc rien de rare***, et on ne peut certainement pas, sur cette base, rejeter l'hypothèse nulle H_0 !

➔ Considérons maintenant le tableau *T'* obtenu à partir de *T* en multipliant les effectifs observés par le nombre strictement positif *k* : pour tout *i* et tout *j*, on a

$$n_{ij}' = k \times n_{ij}.$$

Dans le cas particulier suivi, il vient ceci.

	A	B	C	Totaux
1	$k \times 32$	$k \times 11$	$k \times 20$	
2	$k \times 27$	$k \times 8$	$k \times 25$	
Totaux				

On voit que les effectifs marginaux et l'effectif total seront alors multipliés par *k* :

$$n_{\bullet j}' = k \times n_{\bullet j}\ ;\ n_{i\bullet}' = k \times n_{i\bullet}\ ;\ N' = k \times N.$$

	A	B	C	Totaux
1	$k \times 32$	$k \times 11$	$k \times 20$	$k \times 63$
2	$k \times 27$	$k \times 8$	$k \times 25$	$k \times 60$
Totaux	$k \times 59$	$k \times 19$	$k \times 45$	$k \times 123$

Il en sera de même des effectifs théoriques c_{ij}' puisque

$$c_{ij}' = \frac{n_{\bullet j}' \times n_{i\bullet}'}{N'} = \frac{(k \times n_{\bullet j}) \times (k \times n_{i\bullet})}{k \times N} = k \times \frac{n_{\bullet j} \times n_{i\bullet}}{N} = k \times c_{ij}.$$

Le cas particulier suivi se conforme à cette règle.

	A	B	C	Totaux
1	$k \times 30{,}22$	$k \times 9{,}73$	$k \times 23{,}05$	$k \times 63$
2	$k \times 28{,}78$	$k \times 9{,}27$	$k \times 21{,}95$	$k \times 60$
Totaux	$k \times 59$	$k \times 19$	$k \times 45$	$k \times 123$

Il résulte de là que la contribution de chaque cellule au calcul du χ^2 est multipliée par k :

$$\frac{(n_{ij}' - c_{ij}')^2}{c_{ij}'} = \frac{(kn_{ij} - kc_{ij})^2}{kc_{ij}} = \frac{k^2(n_{ij} - c_{ij})^2}{kc_{ij}} = k\,\frac{(n_{ij} - c_{ij})^2}{c_{ij}}.$$

Si on note kT le tableau T', on peut ainsi écrire : $\chi^2(kT) = k\chi^2(T)$. En d'autres termes, si on multiplie tous les effectifs par k, le χ^2 est multiplié par k : χ^2 est ce qu'on nomme une fonction homogène de degré 1. C'est là un fait essentiel pour élucider le passage du livre à élucider – la sensibilité du χ^2 aux effectifs.

➔ Vérifions tout cela à l'aide du logiciel en ligne déjà utilisé. On prend d'abord $k = 5$, ce qui donne le tableau $T' = 5T$ suivant.

	A	B	C	Totaux
1	**160**	**55**	**100**	
2	**135**	**40**	**125**	
Totaux				

On examinera la réponse du logiciel ci-après.

r × *c* Contingency Table: Results

The results of a contingency table X^2 statistical test performed at 15:03 on 31-MAR-2008

```
data: contingency table

       A       B       C

1     160      55     100     315
2     135      40     125     300

      295      95     225     615



expected: contingency table

        A           B           C

1     151.        48.7        115.
2     144.        46.3        110.
```

chi-square = 6.90
degrees of freedom = 2
probability = 0.032

Cette fois, l'événement observé a, sous l'hypothèse d'indépendance, une probabilité de 3,2 % environ. On observe surtout que la valeur de χ^2 est bien multipliée par 5 : on avait $\chi^2 \approx 1{,}38$, on a maintenant $\chi^2 \approx 6{,}90 = 5 \times 1{,}38$.

➔ Prenons maintenant $k = 8$; le tableau $T' = 8T$, puis la réponse du logiciel, suivent.

	A	B	C	Totaux
1	**256**	**88**	**160**	
2	**216**	**64**	**200**	
Totaux				

chi-square = 11.0
degrees of freedom = 2
probability = 0.004

La valeur de χ^2 est multiplié (approximativement) par 8 ; la probabilité est inférieure à 0,5 %.

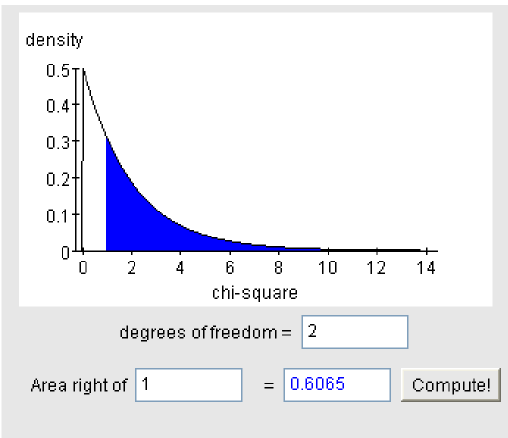
➔ Si l'on prend $k = 20$, on a le tableau ci-après.

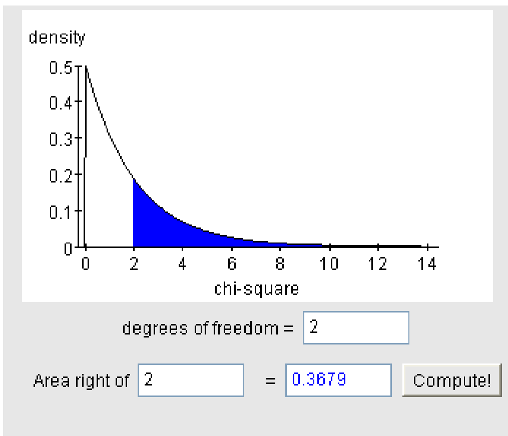
	A	B	C	Totaux
1	**640**	**220**	**400**	
2	**540**	**160**	**500**	
Totaux				

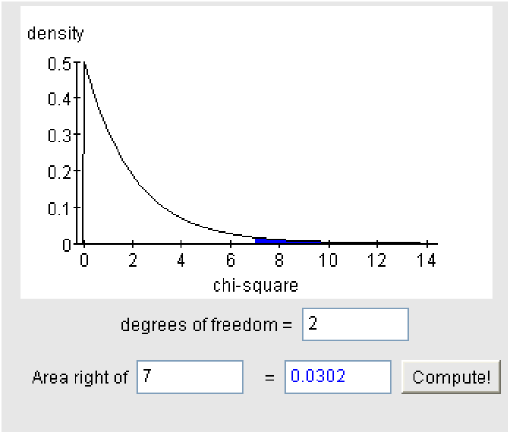
Les résultats suivent : comme on le verra, on a bien $\chi^2 \approx 27{,}6 = 1{,}38 \times 20$. Cette fois, la probabilité affichée est nulle.

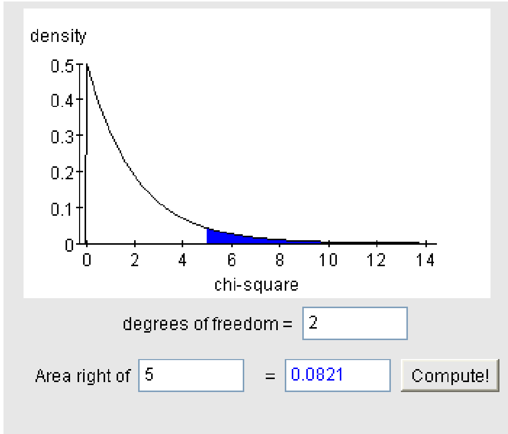
chi-square = 27.6
degrees of freedom = 2
probability = 0.000

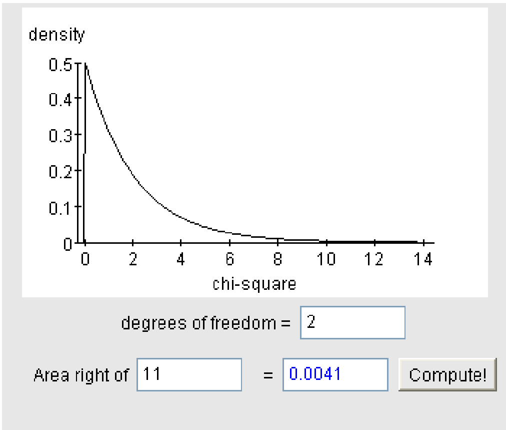
➔ Pour visualiser le phénomène précédent, on recourt maintenant à un logiciel déjà utilisé dans la leçon 2 (http://www.stat.sc.edu/~west/applets/chisqdemo.html).











Ce qu'il convient de noter c'est que l'aire représentant la probabilité d'observer un χ^2 supérieur ou égal à une certaine valeur est calculé chaque fois ***pour la même courbe***, celle représentant la loi du χ^2 à (ici) ***deux*** degrés de liberté.

➔ On reprend ici le tableau étudié dans la leçon 2.

	Blond	Brun	Noir	Roux	Totaux
Bleu	1768	807	189	47	2811
Gris-vert	946	1387	746	53	3132
Brun	115	438	288	16	857
Totaux	2829	2632	1223	116	6800

On a vu que l'on avait ici $\chi^2 \approx 1070$ et que $P(\{\chi^2 \geq 1070\ \} \ /\ H_0) \approx 0$. Considérons le tableau obtenu en divisant les effectifs par 10 (et en arrondissant les quotients obtenus, en rendant compatibles effectifs « observés » et effectifs marginaux).

	Blond	Brun	Noir	Roux	Totaux
Bleu	177	80	19	5	281
Gris-vert	95	139	74	5	313
Brun	11	44	29	2	86
Totaux	283	263	122	12	680

La valeur de χ^2 doit être ici d'environ 107, valeur pour laquelle la probabilité correspondante devrait demeurer nulle. C'est ce que confirme le logiciel utilisé plus haut (voir ci-après).

chi-square = 109.
degrees of freedom = 6
probability = 0.000

On peut alors diviser encore par 10 (c'est-à-dire diviser par 100 le tableau initial).

	Blond	Brun	Noir	Roux	Totaux
Bleu	17	8	2	1	28
Gris-vert	9	14	7	1	31
Brun	2	4	3	0	9
Totaux	28	26	12	2	68

Les résultats affichés par le logiciel utilisé (voir ci-après) montrent que la valeur de χ^2 est à peu près divisée par 10 ; mais cette fois, la probabilité d'observer un tel tableau sous l'hypothèse d'indépendance n'est plus négligeable : elle est un peu supérieure à 15 %.

chi-square = 9.40
degrees of freedom = 6
probability = 0.152

e) Le travail d'excription effectué gagne à être complété d'un travail de ***réinscription***, qui pourrait prendre, ici, la forme du texte ci-après, rédigé par *y*.

Remarque.

♦ Soit T un tableau de contingence à m lignes et p colonnes, dont l'effectif observé correspondant à la ligne $i = 1, 2, \ldots, m$ et à la colonne $j = 1, 2, \ldots, p$ est noté n_{ij}. On note $n_{i\bullet}$ l'effectif marginale de la ligne i et $n_{\bullet j}$ l'effectif marginale de la colonne j. On a :

$$n_{i\bullet} = \sum_{j=1}^{p} n_{ij} \text{ (pour } 1 \leq i \leq m\text{)} \; ; \; n_{\bullet j} = \sum_{i=1}^{m} n_{ij} \text{ (pour } 1 \leq j \leq p\text{)}$$

♦ Considérons le tableau croisé $T' = kT$ obtenu en multipliant les effectifs observés n_{ij} par le nombre $k > 0$, en sorte qu'on a $n_{ij}' = kn_{ij}$, pour tout $i = 1, 2, \ldots, m$ et tout $j = 1, 2, \ldots, p$. On voit que les effectifs marginaux sont eux-mêmes multipliés par k : pour tout i, on a en effet

$$n_{i\bullet}' = \sum_{j=1}^{p} n_{ij}' = \sum_{j=1}^{p} kn_{ij} = k \sum_{j=1}^{p} n_{ij} = k\, n_{i\bullet}$$

et pour tout j, de même, $n_{\bullet j}' = \sum_{i=1}^{m} n_{ij}' = \sum_{i=1}^{m} kn_{ij} = k \sum_{i=1}^{m} n_{ij} = k\, n_{\bullet j}$.

♦ Il en va pareillement pour l'effectif total N : on a par exemple

$$N' = \sum_{i=1}^{m} n_{i\bullet}' = \sum_{i=1}^{m} kn_{i\bullet} = k \sum_{i=1}^{m} n_{i\bullet} = kN.$$

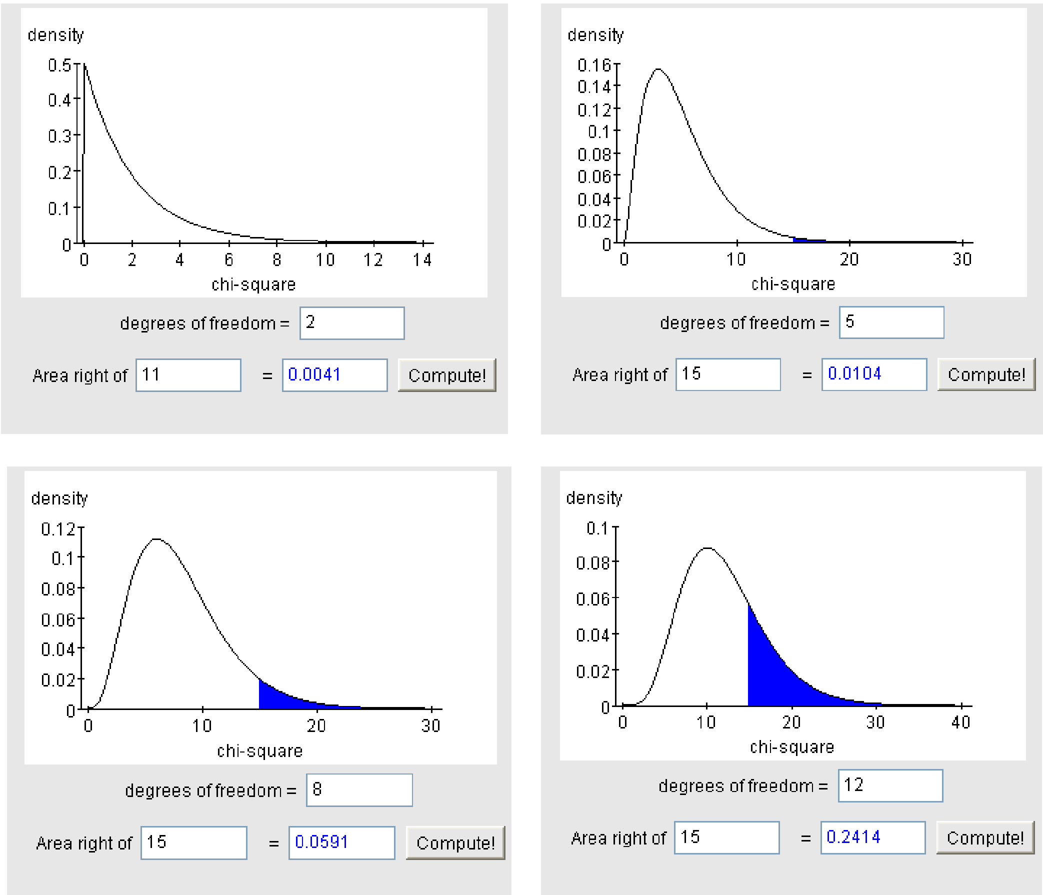
♦ Il résulte de là que les effectifs théoriques c_{ij} sont aussi multipliés par k : pour tout i et tout j on a en effet

$$c_{ij}' = \frac{n_{\bullet j}' \times n_{i\bullet}'}{N'} = \frac{kn_{\bullet j} \times kn_{i\bullet}}{kN} = k\,\frac{n_{\bullet j} \times n_{i\bullet}}{N} = kc_{ij}.$$

♦ En conséquence, la contribution de chacune des cellules du tableau au χ^2 est également multipliée par k, puisque l'on a : $\frac{(n_{ij}' - c_{ij}')^2}{c_{ij}'} = \frac{(kn_{ij} - kc_{ij})^2}{kc_{ij}} = \frac{k^2(n_{ij} - c_{ij})^2}{kc_{ij}} = k\,\frac{(n_{ij} - c_{ij})^2}{c_{ij}}$.

♦ On en conclut que le χ^2 est multiplié par k, ce qu'on peut écrire : $\chi^2(kT) = k\chi^2(T)$. Le χ^2 est donc une fonction homogène de degré 1 des effectifs observés n_{ij}. Cela explique que, lorsqu'on passe d'un tableau T pour lequel l'hypothèse H_0 d'indépendance des caractères croisés ne peut être rejetée à un seuil $\varepsilon > 0$ donné à un tableau kT avec k « assez grand », l'hypothèse H_0 peut être rejetée au seuil indiqué : pour le dire autrement, le test du χ^2 devient rapidement inutile quand les effectifs croissent.

f) Parmi les points qu'il resterait à élucider figure à l'évidence la notion de nombre de ***degrés de liberté***. Quelle relation cette expression entretient-elle avec la structure d'un tableau de contingence T ? Pourquoi ce nombre est-il donné par le produit $(m - 1)(p - 1)$, où m est le nombre de lignes de T et p est le nombre de colonnes de T ? Sur ces points, y invite les $x \in X$ à mener ***leur propre enquête***. On notera ici seulement ceci : pour une valeur de χ^2 donnée (on a pris ci-dessous $\chi^2 = 15$), la probabilité de l'événement ainsi observé, soit $\{ \chi^2 \geq 15 \}$, ***dépend*** clairement du nombre de degrés de liberté, comme on le constatera ci-après.



1. Excrire : *suite*

a) On revient ici au travail engagé autour de la question $Q^\#$: Comment analyser des données ? Plus précisément, on revient au travail amorcé sur le chapitre I (« Repérer les questions pertinentes ») du livre de Philippe Cibois intitulé *Les méthodes d'analyse d'enquêtes*.

➔ La suite de la première partie du chapitre, pp. 9-11, comporte des développements des plus intéressants, mais ne semble pas receler de difficulté majeure. On en vient donc à la section I, intitulée « Première étape : les préalables ». Les mêmes remarques que précédemment peuvent être faites : le texte est instructif sans être trop difficile. Nous nous arrêterons toutefois à la mention du premier tableau croisé, que voici.

Un premier tableau croisé va nous permettre de voir de premiers résultats et de mettre au point un outil qui nous servira dans la suite : un indicateur de la force de liaison entre modalités (ou entre questions). Nous effectuons donc le tri croisé entre le sexe et la variable d'intérêt, le type de collège.

```
Croisement de la question 17A type
d'écoles avec la question V02 sexe
Le Khi-deux du tableau est de    3.8
Degré liberté =    1 Prob.= 0.047 **
```

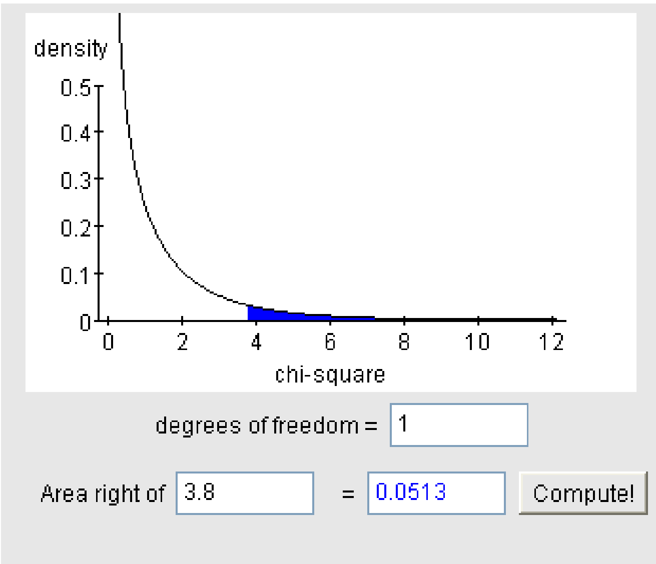
Il s'agit d'un tableau à 2 lignes et 2 colonnes et donc à un degré de liberté puisqu'en fixant l'effectif d'une case, toutes les autres se déduisent des marges.

Le khi-deux est significatif au seuil de 5 % ($p < 0{,}05$ codé souvent avec deux étoiles).

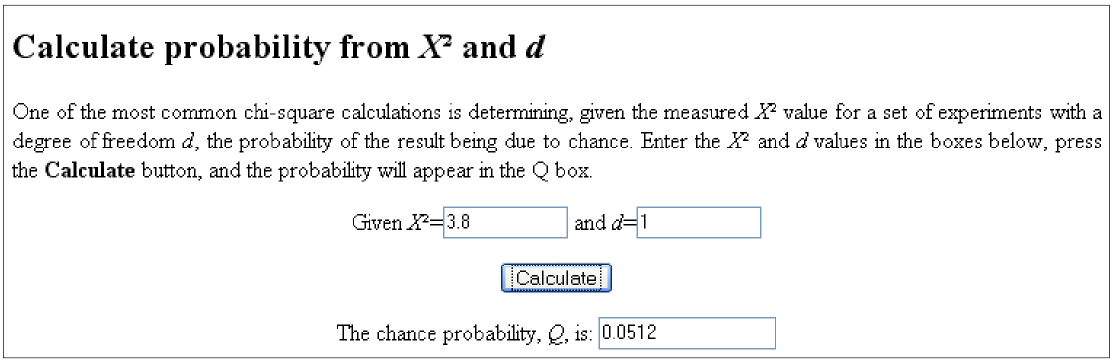
➔ Le tableau mentionné page 14 croise deux variables dichotomiques, le sexe (masculin, féminin) et le type de collège (« École nouvelle », « Collège de bonne réputation ») : c'est donc un tableau 2×2, que nous schématiserons ainsi.

	EN	CBR	Totaux
Masculin	n_{11}	n_{12}	
Féminin	n_{21}	n_{22}	
Totaux			

Le nombre de degrés de liberté est 1 ; comme $\chi^2 = 3{,}8$, on a ceci.



La probabilité indiquée ne correspond pas à celle affichée dans le passage ci-dessus : elle était inférieure à 5 %, elle est maintenant trouvée supérieure (très légèrement) à 5 %. La vérification avec un autre logiciel, également utilisé dans la leçon 2 (http://www.fourmilab.ch/rpkp/experiments/analysis/chiCalc.html), confirme ce dernier résultat.



Calculate probability from X^2 and d

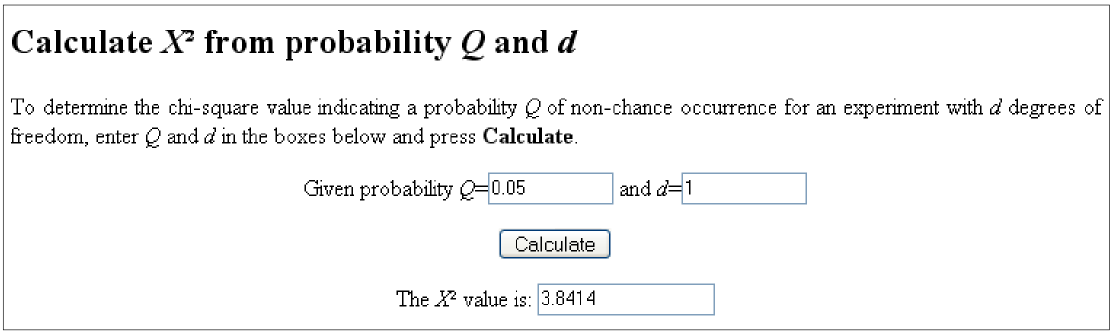
One of the most common chi-square calculations is determining, given the measured X^2 value for a set of experiments with a degree of freedom d, the probability of the result being due to chance. Enter the X^2 and d values in the boxes below, press the **Calculate** button, and the probability will appear in the Q box.

Given X^2= 3.8 and d= 1

Calculate

The chance probability, Q, is: 0.0512

Pour obtenir une probabilité de 5 %, il convient, selon le même logiciel, que l'on ait $\chi^2 > 3{,}84$.



Calculate X^2 from probability Q and d

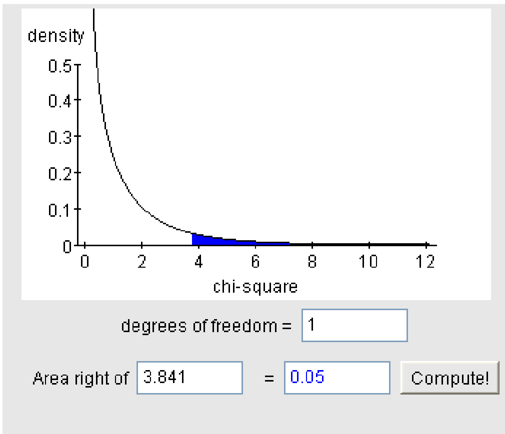
To determine the chi-square value indicating a probability Q of non-chance occurrence for an experiment with d degrees of freedom, enter Q and d in the boxes below and press **Calculate**.

Given probability Q= 0.05 and d= 1

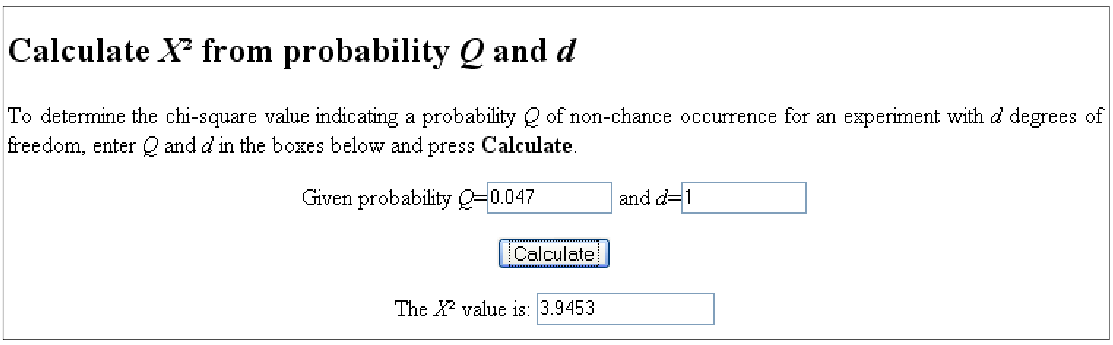
Calculate

The X^2 value is: 3.8414

On peut contrôler ce résultat avec le logique graphique employé plus haut.



Bien entendu, rien n'interdit *a priori* de penser que la valeur affichée par le logiciel Trideux, à savoir 3,8, soit une valeur arrondie (de 3,84…). Mais pour obtenir une probabilité de 0,047, il faut, d'après les logiciels consultés ici, un χ^2 supérieur à 3,94 (voir ci-après).



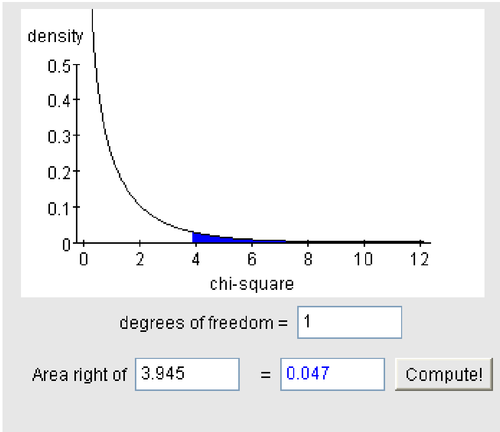
Calculate X^2 from probability Q and d

To determine the chi-square value indicating a probability Q of non-chance occurrence for an experiment with d degrees of freedom, enter Q and d in the boxes below and press **Calculate**.

Given probability Q= 0.047 and d= 1

Calculate

The X^2 value is: 3.9453



➔ En fait, les effectifs observés du tableau précédent figure sur le tableau de la page 16 ; on les reprend ici.

	EN	CBR	Totaux
Masculin	38	206	244
Féminin	60	208	268
Totaux	98	414	512

Si l'on confie le calcul du χ^2 au logiciel déjà utilisé pour ce genre de calcul, on obtient ceci.

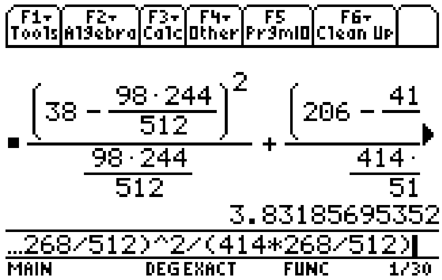
r × *c* Contingency Table: Results

The results of a contingency table X^2 statistical test performed at 06:36 on 1-APR-2008

```
data: contingency table

       A       B

1      38      206      244
2      60      208      268

       98      414      512


expected: contingency table

        A           B

1     46.7        197.
2     51.3        217.
```

chi-square = 3.83
degrees of freedom = 1
probability = 0.050

Si l'on reprend la valeur 3,83, on obtient pour probabilité, selon les deux autres logiciels utilisés la valeur 0,0503. Pour aller plus loin dans le contrôle, on peut calculer plus précisément la valeur de χ^2 : à l'aide d'une calculatrice scientifique, on obtient ce que montre la copie d'écran ci-après.



Pour la valeur 3,831856, le logiciel Excel retourne la probabilité 0,05028723 : on a reproduit ci-dessous les probabilités correspondant à certaines valeurs voisines.

$\chi^2 \geq \ldots$	Probabilité
3,83	0,05034294
3,831856	0,05028723
3,84	0,05004353
3,85	**0,049746**

On retrouve des conclusions déjà formulées : il faut que le χ^2 soit au-dessus de 3,84 pour que la probabilité descende au-dessous de 5 %. La fonction inverse proposée par Excel (KHIDEUX.INVERSE) donne pour valeur de χ^2 correspondant à la probabilité 0,05 le nombre 3,84145534.

➔ Les anomalies constatées appellent à méditer deux conclusions. Tout d'abord, même les logiciels et les livres élaborés par les personnes des plus compétentes peuvent comporter des erreurs : la chose va sans dire, mais elle va mieux en le disant. Ensuite, et surtout ici, une erreur apparente, très légère, conduirait dans un cas à rejeter l'hypothèse nulle au seuil de 5 %, dans l'autre… peut-être pas – du moins, si l'on est très formel ! Il faut donc prendre ses distances par rapport au formalisme et aux recettes qui sont mises à leur service. Décider ou non de rejeter – jusqu'à plus ample informé – l'hypothèse d'indépendance ne doit pas tenir à des poussières de décimales.

b) Avançons dans le texte examiné : c'est une nouvelle difficulté qui va s'élever devant nous et qu'il nous faudra donc excrire.

➔ Celui-ci se poursuit par le développement ci-après. On y rencontre immédiatement un nouveau venu : le « PEM ».

Le PEM du tableau est de 18.6%

Par PEM, il faut entendre Pourcentage de l'écart maximum : il s'agit d'un indicateur d'attraction qui vaut pour l'ensemble du tableau (PEM global) ou pour une case du tableau (PEM local). Nous allons expliciter en premier lieu le PEM local.

➔ Du PEM nous n'apprenons d'abord que le nom – « Pourcentage de l'écart maximum » – et la fonction : ce serait un « indicateur d'attraction » valant tant « pour l'ensemble du tableau (PEM global) » que « pour une case du tableau (PEM local) ». Cela est bien peu ; la suite va être plus parlante : nous la commenterons au fur et à mesure entre crochets. Mais avant cela nous reproduisons le tableau de la page 16 de l'ouvrage.

N= %Ligne Khi2%Attrac	Ecole Nouvelle	Collège BonneRép	Total en ligne
Masc	38 15.6 1.6 -18.6*	206 84.4 0.4 18.6*	244 100 2.0 47.7
Fémi	60 22.4 1.5 18.6*	208 77.6 0.3 18.6*	268 100 1.8 52.3
Total en colonne	98 19.1 3.1	414 80.9 0.7	512 100 3.8 100

Voici alors le texte commenté annoncé.

Dans le tableau p. 16, on trouve quatre nombres dans chaque case (et leur somme en marge).

– L'effectif (N =) : pour la case « Féminin en École nouvelle », il est de 60 individus.

– Le pourcentage en ligne (% Ligne) : sur 268 élèves de sexe féminin, les 60 en école nouvelle représentent 22,4 % [on a $\frac{60}{268}$ = 0,22388… ≈ 22,388 % ≈ 22,4 %] du total (soit plus que 19,1 %, le pourcentage toutes lignes confondues, ce qui indique une attraction).

– La contribution au khi-deux qui est égale à l'effectif en écart à l'indépendance au carré divisé par l'effectif théorique.

Ici l'effectif théorique (produit des marges par le total) est de 98 × 268 / 512 = 51,30 [on a $\frac{98 \times 268}{512}$ = 51,296875 ≈ 51,30]. L'écart à l'indépendance est de (observé – théorique) 60 – 51,30 = 8,70 [on a 60 – $\frac{98 \times 268}{512}$ = 8,703125 ≈ 8,7]. La contribution au khi-deux est de $8{,}70^2/51{,}30$ = 1,5 [on a $\frac{8{,}703125^2}{51{,}296875}$ = 1,4765886765… ≈ 1,48]

– Le PEM, Pourcentage de l'écart maximum (% Attrac). On a noté que pour cette case, l'écart à l'indépendance est 8,70 individus [exactement, 8,703125]. Si la liaison entre sexe féminin et école nouvelle était à son maximum, les 268 filles ne pourraient pas être à l'école nouvelle (dont l'effectif n'est que de 98 individus) mais les 98 élèves de l'école nouvelle pourraient être de sexe féminin. Donc 98 est le maximum de la case et l'écart à l'indépendance dans le cas de ce maximum serait de (maximum – théorique) 98 – 51,30 = 46,70 [plus exactement 46,703125].

Comparons l'écart observé 8,70 à l'écart dans le cas du maximum 46,70 ce qui nous donne une proportion de 8,70 / 46,70 = 0,186 ou 18,6 % en pourcentage [plus exactement, $\frac{8{,}703125}{46{,}703125}$ = 0,1863499… ≈ 18,6 %]. Cette valeur est suivie d'une étoile sur le tableau pour signaler qu'elle est issue d'un tableau croisé significatif [*sic*].

➔ Pour contrôler la notion de PEM, on peut se proposer de calculer les PEM des autres cellules. Pour la cellule (1, 1), l'effectif théorique est

$$c_{11} = \frac{98 \times 244}{512} = \frac{2989}{64} = 46{,}703125.$$

Le PEM est-il bien égal, ici, à

$$\frac{n_{11}-c_{11}}{98-c_{11}}=\frac{38-\frac{2989}{64}}{98-\frac{2989}{64}}=-0{,}16966\ldots\approx -17\ \%\ ?$$

Le tableau de l'auteur donne –18,6 (%). En fait, ici, l'écart

$$n_{11}-c_{11}=38-46{,}703125$$

est négatif ; l'écart « maximum » (en valeur absolue) semble devoir être obtenu quand on remplace n_{11} par 0 : il vaut alors – 46,703125 et le pourcentage de l'écart « maximum » s'écrit

$$\frac{n_{11}-c_{11}}{0-c_{11}}=\frac{38-\frac{2989}{64}}{-\frac{2989}{64}}=0{,}1863499\ldots\approx 18{,}6\ \%.$$

Il y a cependant un hic : dans le tableau proposé par l'auteur, la cellule (1, 1) comporte, non le nombre 18,6 mais son opposé, – 18,6. Laissons cette difficulté en attente et passons à la suite du tableau.

➔ La cellule (1, 2) a pour effectif observé $n_{12}=206$. Son effectif théorique est

$$c_{12}=\frac{414\times 244}{512}=\frac{12627}{64}=197{,}296875.$$

On se trouve ici dans le cas où l'effectif observé est supérieur à l'effectif théorique ; le PEM est le rapport de l'écart $n_{12}-c_{12}$ à l'écart « maximum » : ici, les 244 garçons pourraient être en CBR (il y a 414 « places »), en sorte que l'écart « maximum » est $244-c_{12}$. Le pourcentage serait alors :

$$\frac{n_{12}-c_{12}}{244-c_{12}}=\frac{206-\frac{12627}{64}}{244-\frac{12627}{64}}=0{,}1863499\ldots\approx 18{,}6\ \%.$$

➔ Passons à la cellule (2, 2). L'effectif observé est $n_{22}=208$; l'effectif théorique vaut

$$c_{22}=\frac{414\times 268}{512}=\frac{13869}{64}=216{,}703125.$$

En ce cas l'écart $n_{22}-c_{22}$ est négatif. Il est « maximum » en valeur absolue quand on remplace n_{22} par 0, ce qui donnerait donc pour PEM

$$\frac{n_{22}-c_{22}}{0-c_{12}}=\frac{208-\frac{13869}{64}}{-\frac{13869}{64}}=0{,}04016\ldots$$

Cela ne correspond pas au nombre affiché dans le tableau : 18,6. En fait, ici, on ne peut remplacer n_{22} par 0 : le minimum de n_{22} est obtenu si le maximum possible des 268 filles vont en EN ; or ce maximum est de 98, et le minimum de filles en CBR est donc 268 – 98 = 170. Le PEM vaut donc :

$$\frac{n_{22} - c_{22}}{170 - c_{12}} = \frac{208 - \frac{13869}{64}}{170 - \frac{13869}{64}} = 0{,}1863499\ldots \approx 18{,}6\ \%.$$

ce qui concorde avec l'indication donnée par le tableau de la page 16 dans la cellule (2, 2).

➔ Il reste à élucider la question du signe du PEM de la cellule (1, 1). Le passage examiné comporte une note de bas de page qui renvoie le lecteur à la référence suivante :

Philippe Cibois, Le PEM, pourcentage de l'écart maximum : un indice de liaison entre modalités d'un tableau de contingence, *Bulletin de méthodologie sociologique*, 1993, n° 40, p. 43-63.

En recherchant une version en ligne de ce texte, on tombe sur le site personnel de Philippe Cibois (http://pagesperso-orange.fr/cibois/Text.html), où, en effet, on trouve bien ce texte (on le trouvera aussi à l'adresse http://www.modalisa.com/pdf/CiboisPEM.pdf), mais où on trouve aussi une liste des « erreurs à corriger dans *Les méthodes d'analyse d'enquêtes* ». À la date du 2 mars 2008, la seule qui concerne les passages examinés jusqu'ici est la suivante : « p. 16 : case "Collège Bonne réputation" – "Féminin", lire – 18.6 pour le PEM (manque le signe moins) » En d'autres termes, l'accord noté pour la cellule (2, 2) doit être remis en cause !

➔ L'examen du texte indiqué par l'auteur permet de « clarifier » le cas du signe lorsque l'effectif observé est inférieur à l'effectif théorique, ce qui se produit dans les cellules (1, 1) et (2,2). Nous renvoyons le lecteur intéressé à l'annexe de l'article en question (pp. 12-15). On y lira notamment cette décision de l'auteur quant à la définition du PEM qui explique notre double « désaccord » avec le tableau (rectifié) de l'auteur : dans le cas où l'écart $n_{ij} - c_{ij}$ est négatif, « EMAX [soit MIN – c_{ij}, où est le minimum remplaçant n_{ij} ainsi qu'on l'a vu] et ECART [soit $n_{ij} - c_{ij}$] sont négatifs le pourcentage [soit PEM = ECART/EMAX] est positif et par convention on inverse son signe pour exprimer le sens de la liaison : PEM = – PEM »…

c) Nous laisserons là le texte de Philippe Cibois. Ce qui précède montre par contraste que les lectures usuelles de textes – « didactiques » ou non – ne sont en règle générale guère excriptrices, mais qu'elles restent « en surface ». Le travail d'excription, en revanche, appelle de la part de qui entend le réaliser de façon idoine ***de fort nombreuses prises d'initiative*** : s'il concourt, comme on l'a imaginé ici, à avancer en un certain PER, il suppose lui-même une multiplicité de PER ***auxiliaires***, souvent des plus inattendus : une enquête est, par nature, une aventure démultipliée.

2. Excrire : généralisation

a) Le travail d'excription a été présenté jusqu'ici comme un travail sur des ***textes***, ce dont la notion de « lecture excriptrice » témoigne. En force domaines, en effet, les textes sont un lieu privilégié d'inscription de praxéologies – personnelles ou institutionnelles –, qu'il s'agisse de textes mus clairement par une intention didactique (manuels, traités, etc.) ou bien de simples chroniques, de narrations, de comptes rendus d'observation, d'entretiens, voire de textes de fiction, etc. Or, de la même façon qu'on peut lire un texte produit par une personne ou une institution, on peut « lire » cette institution ou cette personne dans le but d'excrire son équipement praxéologique relatif à tel ou tel domaine d'activité, parce que la connaissance de celui-ci apparaît utile à l'avancement de tel projet personnel ou institutionnel. On retrouve alors dans l'activité d'excription praxéologique l'exigence dont on a tenté de donner une idée dans tout ce qui précède – celle de la ténacité, de la persévérance excriptrices nécessaires pour ne pas demeurer à la surface des choses. La brièveté de ce cours nous contraint à ne pas aller beaucoup plus loin sur cet immense sujet.

b) Nous illustrerons le thème de l'observation praxéologique et de l'excription personnelle ou institutionnelle qu'elle doit viser en nous arrêtant ici sur un geste d'apparence mineur : l'exploration à l'aide d'un différenciateur sémantique de certains éléments technologiques et théoriques personnels ou institutionnels à propos de telle ou telle réalité sociale. On commencera en présentant les résultats d'un tel geste de prise d'information praxéologique sur un groupe de 22 étudiants de licence de sciences de l'éducation en formation à Lambesc.

➔ Voici d'abord les scores, totalisés sur l'ensemble des juges, obtenus par « Internet » sur les 15 échelles qui étaient proposées.

ℓ/L	S/P	F/C	T/B	PU/U	I/A	M/B	f/F	A/C	S/O	**L/B**	P/A	I/S	A/N	F/D
54	50	54	80	108	57	83	93	74	80	**68**	94	45	33	45

On voit que le score du mot « Internet » sur l'échelle laid/beau, 68, se situe presque exactement à mi-chemin entre le minimum théorique (0) et le maximum théorique ($22 \times 6 = 132$). À juger selon ce critère numérique, « Internet » ne serait ni laid ni beau pour notre groupe de juges, même si l'un d'eux lui attribue le score maximal, 6. Les échelles sur lesquelles les scores de « Internet » sont les plus extrêmes sont mises en évidence ci-après.

ℓ/L	S/P	F/C	T/B	**PU/U**	I/A	M/B	f/F	A/C	S/O	L/B	P/A	I/S	**A/N**	F/D
54	50	54	80	**108**	57	83	93	74	80	68	94	45	**33**	45

Le score 108 sur l'échelle PU/U, c'est-à-dire sur l'échelle peu utile/utile, correspond à 81,8 % du maximum. (Sur l'ensemble des mots et des échelles, le plus haut score total est 123, soit 93,2 % du maximum.) À l'évidence, « Internet » est ***fort utile*** (le plus bas score attribué à ce mot est 3, la « moyenne », qui est obtenu deux fois). Le score de 33 obtenu sur l'échelle A/N, soit l'échelle artificiel/naturel, représente 25 % du maximum : en lisant l'échelle dans l'autre sens (naturel/artificiel), on arrive à un score total de 99, soit (bien sûr) 75 % du maximum. C'est le plus bas score total obtenu sur l'ensemble des mots : « Internet » est ainsi ***fortement artificiel***. Les échelles sur lesquelles « Internet » obtient un score total supérieur à 65 % du maximum ou inférieur à 35 % du maximum, c'est-à-dire supérieur à 85,8 points ou inférieur à 46,2, sont mises en évidence ci-après.

ℓ/L	S/P	F/C	T/B	PU/U	I/A	M/B	**f/F**	A/C	S/O	L/B	**P/A**	**I/S**	A/N	**F/D**
54	50	54	80	108	57	83	**93**	74	80	68	**94**	**45**	33	**45**

Il s'agit, de gauche à droite, des échelles faible/fort, passif/actif, instable/stable et facile/difficile : « Internet » est donc jugé dans l'ensemble plutôt ***fort***, ***actif***, ***instable*** et… ***facile***. En vérité, les deux scores de 45 dissimulent des situations un peu différentes. Sur le fait que « Internet » serait plutôt ***instable***, il y a un accord assez large du groupe des juges, comme le montre le tableau suivant.

Score	0	1	2	3	4	5	6
Effectif	3	4	5	9	1	0	0

En revanche, les jugements sur le fait qu'« Internet » serait ou non ***difficile*** sont davantage contrastés, ainsi qu'en témoigne le tableau suivant.

Score	0	1	2	3	4	5	6
Effectif	3	6	5	4	3	1	0

➔ On peut comparer les scores obtenus par « Internet » et « Web ». Sur le tableau suivant, la deuxième ligne rappelle les scores totaux de « Internet » ; la troisième présente ceux relatifs à « Web ».

ℓ/L	S/P	F/C	T/B	PU/U	I/A	M/B	f/F	A/C	S/O	L/B	P/A	I/S	A/N	F/D
54	50	54	80	108	57	83	93	74	80	68	94	45	33	45
74	55	52	76	105	61	79	86	78	88	71	83	55	34	55

On voit que les différences ne semblent pas profondes : « Web » apparaît par exemple un peu plus ***lourd*** (ℓ/L, léger/lourd) que « Internet » mais moins ***actif*** (P/A, passif/actif). Cela noté, on peut se demander si « Internet » et « Web » ont la particularité d'avoir seuls le « profil » que nous avons sommairement décrit jusqu'ici. Voici donc le tableau précédent complété par des lignes relatives aux autres mots proposés aux 22 juges, à savoir, après « Internet » et « Web », et par ordre croissant des numéros des lignes du tableau : « Statistique », « Didactique », « Université », « Éducation », « Lambesc ».

ℓ/L	S/P	F/C	T/B	PU/U	I/A	M/B	f/F	A/C	S/O	L/B	P/A	I/S	A/N	F/D
54	50	54	80	108	57	83	93	74	80	68	94	45	33	45
74	55	52	76	105	61	79	86	78	88	71	83	55	34	55
85	73	38	60	95	47	79	80	76	107	56	85	66	45	80
85	100	80	84	106	73	100	102	88	91	78	88	79	71	83
95	108	75	112	112	67	100	107	79	96	88	105	74	73	98
94	115	95	91	123	69	109	105	91	85	102	114	78	81	94
62	109	84	98	117	77	108	104	85	88	101	108	84	79	84

On observe ici, dans plusieurs colonnes, un changement qui se produit lorsqu'on passe du bloc des trois premières lignes à celui des quatre dernières. Ainsi, sur l'échelle S/P, superficiel/profond, on découvre que « Internet » et « Web » sont plutôt ***superficiels***, que c'est un peu moins vrai pour « Statistique », et que « Didactique », « Université », « Éducation » et « Lambesc » sont regardés bien davantage comme ***profonds***. Un schéma analogue s'observe à propos de l'échelle F/C, froid/chaud : « Statistique » surtout, suivis par « Internet » et « Web » sont ***froids***, tandis que « Didactique », « Université », « Éducation » et « Lambesc » sont plutôt chauds. Sur l'échelle T/B, terne/brillant, une étoile se détache : « Université » ; lui fait pendant un astre relativement terne, « Statistique ». Sur l'échelle PU/U, tous les mots sont rangés franchement du côté de l'***utile***, y compris « Statistique », avec un pic, « Éducation », dont le score, qui représente 93,2 % du maximum, est le maximum maximorum. La moindre appréciation de « Statistique » s'exprime encore sur l'échelle suivante, I/A, irritant/apaisant, où il est le moins bien noté : ici, en vérité, aucun mot n'a de score impressionnant puisque le maximum est atteint par « Lambesc », dont le score ne représente pourtant que 58,3 % du score maximal. On retrouve le contraste entre les trois premières lignes et les quatre suivantes avec l'échelle M/B, mauvais/bon : « Internet », « Web » et « Statistique » ne sont pas mauvais, mais « Didactique », « Université », « Éducation » et « Lambesc » sont meilleurs. On laissera le lecteur poursuivre ce rapide examen, sachant que les colonnes suivantes du tableau correspondent successivement aux échelles faible/fort (f/F), abstrait/concret (A/C), spontané/organisé (S/O), laid/beau (L/B), passif/actif (P/A), instable/stable (I/S), artificiel/naturel (A/N).

➔ L'information recueillie et analysée (sommairement !) nous éclaire, de façon bien sûr conjecturale, sur le rapport des « sujets » examinés à travers un différenciateur sémantique à un certains nombre d'objets : il est remarquable, ainsi, qu'il y ait un bon accord autour de ***l'utilité*** qu'auraient les réalités évoquées par les mots retenus pour l'examen réalisé. On peut même imaginer qu'il s'agisse là d'une attitude systématique, engendrée par l'engagement formatif de ce type de « public » : pour cette raison, on peut s'attendre à ce qu'un autre mot, désignant une autre réalité de l'univers de ces « sujets », obtienne un score élevé sur l'échelle peu utile/utile. Mais on ne manquera pas d'observer aussi que l'effort de formation ainsi apparemment consenti va de pair avec ce qu'on peut décrire comme une imparfaite sérénité ! On n'insistera pas, enfin, sur ce que révèle l'analyse ébauchée à propos de la « théorie » des sujets observés touchant la ***statistique*** – discipline froide, plutôt terne, irritante aussi, pas particulièrement mauvaise sans doute mais très organisée, dénuée de beauté, et pour finir artificielle…

c) On en vient ici à présenter les résultats de la prise d'information réalisée auprès de 10 étudiants de première année du master recherche de sciences de l'éducation en formation à Aix.

➔ Rappelons d'abord quelles étaient les échelles, au nombre de 10 : superficiel/profond (S/P), froid/chaud (F/C), peu utile/utile (PU/U), irritant/apaisant (I/A), peu fiable/fiable (P/F), abstrait/concret (A/C), inorganisé/organisé (I/O), laid/beau (L/B), passif/actif (P/A), facile/difficile (F/D).

➔ Les mots auxquels il fallait attribuer des scores sur ces échelles étaient les suivants (dans l'ordre croissant des lignes du tableau ci-après) : université ; statistique ; éducation ; Internet ; mémoire (le).

➔ Voici le tableau des scores totalisés sur les répondants.

S/P	F/C	PU/U	I/A	P/F	A/C	I/O	L/B	P/A	F/D
45	34	51	31	45	33	32	33	47	40
34	27	47	28	37	38	50	36	31	30
55	48	55	39	39	42	39	47	54	49
33	35	49	35	36	42	36	36	37	23
41	43	48	27	39	37	46	37	50	44

➔ Une première lecture peut consister à classer les mots sur les différentes échelles. Voici d'abord ces classements, sans les scores attribués.

☞ **superficiel/profond :**
Internet ≺ statistique ≺ mémoire ≺ université ≺ éducation

☞ **froid/chaud :**
statistique ≺ université ≺ Internet ≺ mémoire ≺ éducation

☞ **peu utile/utile :**
statistique ≺ mémoire ≺ Internet ≺ université ≺ éducation

☞ **irritant/apaisant :**

mémoire ≺ statistique ≺ université ≺ Internet ≺ éducation

☞ **peu fiable/fiable :**
Internet ≺ statistique ≺ éducation ~ mémoire ≺ université

☞ **abstrait/concret :**
université ≺ mémoire ≺ statistique ≺ éducation ~ Internet

☞ **inorganisé/organisé :**
université ≺ Internet ≺ éducation ≺ mémoire ≺ statistique

☞ **laid/beau :**
université ≺ statistique ~ Internet ≺ mémoire ≺ éducation

☞ **passif/actif :**
statistique ≺ Internet ≺ université ≺ mémoire ≺ éducation

☞ **facile/difficile :**
Internet ≺ statistique ≺ université ≺ mémoire ≺ éducation

➔ Voici maintenant ces mêmes classements avec les scores totaux obtenus.

☛ **superficiel/profond :**
Internet (33) ≺ statistique (34) ≺ mémoire (41) ≺ université (45) ≺ éducation (55)

☛ **froid/chaud :**
statistique (27) ≺ université (34) ≺ Internet (35) ≺ mémoire (43) ≺ éducation (48)

☛ **peu utile/utile :**
statistique (47) ≺ mémoire (48) ≺ Internet (49) ≺ université (51) ≺ éducation (55)

☛ **irritant/apaisant :**
mémoire (27) ≺ statistique (28) ≺ université (31) ≺ Internet (35) ≺ éducation (39)

☛ **peu fiable/fiable :**
Internet (35) ≺ statistique (37) ≺ éducation (39) ~ mémoire (39) ≺ université (45)

☛ **abstrait/concret :**
université (33) ≺ mémoire (37) ≺ statistique (38) ≺ éducation (42) ~ Internet (42)

☛ **inorganisé/organisé :**
université (32) ≺ Internet (36) ≺ éducation (39) ≺ mémoire (46) ≺ statistique (50)

☛ **laid/beau :**
université (33) ≺ statistique (36) ~ Internet (36) ≺ mémoire (37) ≺ éducation (47)

☛ **passif/actif :**
statistique (31) ≺ Internet (37) ≺ université (47) ≺ mémoire (50) ≺ éducation (54)

☛ **facile/difficile :**
Internet (23) ≺ statistique (30) ≺ université (40) ≺ mémoire (44) ≺ éducation (49)

➔ Un certain nombre de conclusions peuvent être explicitées. Le mot qui a sans doute la meilleure image est « éducation » : il obtient 55/60, soit 91,7 % du maximum, sur deux échelles, S/P et PU/U ; mais il obtient encore 54/60, soit 90 % du maximum, sur l'échelle P/A. Son portrait : ***très profond***, ***très utile***, ***très actif***, mais ***difficile*** aussi (49/60, soit 81,7 %), ***chaud*** (48/60, soit 80 %), de loin le plus ***beau*** de tous les mots proposés (47/60, soit 78,3 %, loin devant ses suivants, dont le score reste inférieur ou égal à 37/60, soit 61,7 %), il est aussi ***le plus concret*** (avec « Internet »), et le ***moins irritant***. Seule ombre au tableau : il n'est que moyennement ***fiable*** et moyennement ***organisé***.

➔ À « éducation » on peut comparer « université ». Ici, le tableau est plus contrasté. Pour ***l'utilité***, il obtient 51/60 et n'est « battu » que par « éducation » (55/60). Mais il est vrai que les cinq mots proposés sont trouvés des plus ***utiles***, le plus faible score, celui de « statistique » étant tout de même de 47/60 : on retrouve ici la forte adhésion d'ensemble aux études entreprises. Avec ce même score exactement, il est jugé ***actif***, même si « éducation » le surclasse à cet égard. « Université » se rapproche encore de « éducation » sur l'échelle S/P : il est ***profond*** « à 75 % » (son score est 45/60) ; il obtient le même score sur l'échelle P/F, peu fiable/fiable, se distinguant en cela de « éducation », jugé moyennement fiable. Bien que ***fiable***, donc, « université » est en revanche nettement ***inorganisé*** – mais il est vrai que « éducation » ne brille guère plus à cet égard. Il est en outre une échelle sur laquelle « université » et « éducation » sont aux antipodes l'un de l'autre : alors que « éducation » est le plus ***beau*** de tous les mots, on l'a vu, « université » est ***le moins beau*** – il est remarquable qu'il soit même plus laid que « statistique » ! Sur deux autres échelles encore « université » a des performances décevantes : il n'est pas très ***chaud*** (même s'il l'est un peu plus que « statistique »), et surtout il est le plus ***abstrait*** de tous les mots. Il occupe une position médiane sur l'échelle I/A (où les scores sont resserrés entre 45 % et 65 % du maximum) ainsi que l'échelle F/D, tout en étant sensiblement plus difficile que « Internet » et surtout que « statistique ».

➔ « Mémoire » est un mot proposé dans une période sensible puisque les « enquêtés » ont le souci de mener à bien le mémoire de première année de master qui leur est demandé. On ne s'étonnera guère, en conséquence, que « mémoire » soit jugé le plus ***irritant*** des mots proposés, au même niveau que « statistique ». De même, s'il est bien entendu ***utile*** (avec un score de 47/60), il ne l'est guère plus que « statistique ». Il est trouvé relativement ***abstrait***, même s'il l'est moins que « université ». Il est assez ***profond***, mais moins que « éducation » ou même que « université ». Plus positivement, il occupe le ***second rang*** derrière « éducation » du point de vue de l'***activité***, de la ***difficulté***, de la ***chaleur***, de la ***beauté***, derrière « université » pour la ***fiabilité***, derrière « statistique » pour ce qui est, enfin, du niveau d'***organisation***.

➔ Les mots « statistique » et « Internet » n'appartiennent pas, à l'évidence, à la même famille que « éducation », « université », « mémoire ». Ce sont, grosso modo, d'assez « mauvais objets » : ainsi sont-ils ***dernier et avant-dernier*** sur les échelles superficiel/profond, peu fiable/fiable, passif/actif, facile/difficile : ***superficiel***, ***peu fiable***, ***passif***, et, coup de pied de l'âne, ***facile***. Le portrait n'est pas flatteur ! Bien entendu, l'un et l'autre sont ***utiles***. Ils ne se séparent guère non plus sur l'échelle laid/beau, où ils se positionnement entre les deux extrêmes – ils sont toutefois nettement moins ***beaux*** que « éducation ». Cela dit, « Internet » l'emporte sur « statistique » sur l'échelle froid/chaud (« statistique » est sensiblement plus froid), sur l'échelle irritant/apaisant (c'est « statistique » qui irrite le plus, avec… « mémoire »), sur l'échelle abstrait/concret (« Internet » est le plus concret de tous les mots proposés). Une seule inversion, qui n'est véritablement à l'avantage d'aucun des deux mots : « statistique » est très ***organisé*** (son score est de 50/60), « Internet » l'est beaucoup moins (36/60). Si le meilleur des mots est « éducation », le pire est donc « statistique ».

d) On réalise maintenant une nouvelle prise d'information, à propos de l'une – et de l'une seulement – des réalités évoquées ci-dessus : ***Internet***.

➔ Voici le questionnaire « Internet et moi » que chacun est invité à remplir.

Internet et moi

Vous vous imaginez dans la situation suivante : un enfant de dix ans environ vous pose les questions ci-après, toutes relatives à Internet.

Pour chacune de ces questions, indiquez quel score (de 0 à 4) mesurant la qualité de votre réponse *possible* vous vous assigneriez *avant* de répondre effectivement à l'enfant.

Question 1. *Internet, qu'est-ce que c'est ?*

0———1———2———3———4

Question 2. *Comment les ordinateurs sont-ils connectés à Internet ?*

0———1———2———3———4

Question 3. *Qu'est-ce qu'un site Internet ?*

0———1———2———3———4

Question 4. *Comment les informations circulent-elles sur Internet ?*

0———1———2———3———4

Question 5. *Peut-on pirater Internet ?*

0———1———2———3———4

Question 6. *Comment fonctionne la messagerie électronique ?*

0———1———2———3———4

➔ Voici maintenant le questionnaire « Internet et moi, *bis* ».

Internet et moi, *bis*

Vous vous imaginez dans la situation suivante : vous avez dû répondre à un enfant de dix ans environ à propos des questions ci-après, toutes relatives à Internet.

Pour chacune de ces questions, précisez le score (de 0 à 4) mesurant ce que vous estimez être votre manque de connaissances pour répondre d'une façon satisfaisante à cet enfant.

Question 1. *Internet, qu'est-ce que c'est ?*

0———1———2———3———4

Question 2. *Comment les ordinateurs sont-ils connectés à Internet ?*

0———1———2———3———4

Question 3. *Qu'est-ce qu'un site Internet ?*

0———1———2———3———4

Question 4. *Comment les informations circulent-elles sur Internet ?*

0———1———2———3———4

Question 5. *Peut-on pirater Internet ?*

0———1———2———3———4

Question 6. *Comment fonctionne la messagerie électronique ?*

0———1———2———3———4

➔ Les données recueillies feront l'objet d'un compte rendu mis en annexe à cette leçon. Les personnes intéressées pourront par ailleurs confronter leurs idées de réponses aux questions ci-dessus à celles données dans le petit livre de Françoise Virieux (illustré par Sébastien Chebret), *Internet : quel drôle de réseau !* Destiné aux 9-12 ans, cet ouvrage a été publié en 2006 aux éditions Le Pommier, dans la collection « Les minipommes ».

3. Se former à l'enquête codisciplinaire

a) Avant de clore ce bref travail sur la notion de PER, il convient de s'arrêter un instant sur une question d'éducation au cœur de notre temps et surtout des temps à venir : celle de la formation – scolaire d'abord, mais aussi extrascolaire – des citoyens à la conception et à la réalisation de PER, dans une position ou une autre, soit comme $x \in X$, soit comme $y \in Y$, pour se référer au schéma de base déjà présenté (voir la leçon 1) et que revoici.

$$[S(X\,;y\,;Q) \Rightarrow \{ R_1^\lozenge, R_2^\lozenge, \ldots, R_n^\lozenge, O_{n+1}, \ldots, O_m \}] \hookrightarrow R^\heartsuit.$$

b) Le parcours d'étude et de recherche dont le bilan se lit dans le schéma précédent est déterminé par le projet d'enquêter sur une question Q, l'intérêt pour apporter une réponse à cette question étant lui-même engendré par un ***projet*** plus vaste, P, qui peut être de nature scientifique, associative, militante, ou toucher à la vie personnelle, au quotidien, aux loisirs, etc. Face à cette ambition éducative d'aujourd'hui, de nombreux obstacles se dressent, dont le premier est sans doute l'habitus pluriséculaire de recevoir ***toutes faites*** des réponses diffusées par des « autorités » supposées à l'adresse de citoyens ***didactiquement passifs*** – *magister dixit*, « le maître l'a dit », selon la célèbre formule scolastique qui coupait court à toute contestation (le maître invoqué étant, en l'espèce, Aristote). La formation à l'enquête est d'abord un antidote à un passé immémorial fait de renoncements et presque tout entier tissé d'obtempérations scandées de rébellions individuelles ou non (le grec *hairésis*, d'où vient le français « hérésie », signifie à l'origine « choix individuel », avant de désigner une « doctrine contraire à l'orthodoxie »), longtemps durement réprimées. Surtout, cette partie de l'éducation moderne pourrait devenir un instrument central de la vie des sociétés démocratiques. Bien entendu, à vouloir enquêter, souvent sous la forme de contre-enquêtes ou de reprises d'enquête, même menées collectivement, chacun prend des ***risques*** au plan de la responsabilité devant soi-même, devant la société ou la microsociété qui aura à connaître et peut-être à juger les fruits de ces PER qui, peut-être, ne s'autorisent d'abord que de la volonté de leurs acteurs.

c) Un autre obstacle est celui de la capacité personnelle ou collective à mobiliser les outils utiles à l'enquête – ce qui, dans le schéma formel rappelé ci-dessus, est noté $R_1^\lozenge, R_2^\lozenge, \ldots, R_n^\lozenge, O_{n+1}, \ldots, O_m$. On a souligné dans ce qui précède la difficulté contre-culturelle du travail d'excription nécessaire pour « arracher » les réponses $R_1^\lozenge, R_2^\lozenge, \ldots, R_n^\lozenge$ aux personnes ou aux institutions qui les portent parfois soit de façon silencieuse, intime, muette, soit de façon volubile mais bruyante, sans garantie de sincérité et d'authenticité.

➔ On soulignera ici un autre obstacle : celui de la légitimité que chacun se reconnaît et qu'on lui reconnaît à accéder aux œuvres utiles O_{n+1}, …, O_m. Notons que, à cet égard, l'éducation scolaire a une dimension schizophrénogène qui nous impose ***du même mouvement*** tout à la fois de fréquenter nombre d'œuvres de la culture et de nous interdire en même temps d'y recourir lorsque la situation l'imposerait. Combien de citoyens, « anciens élèves » pour l'éternité, assument sans faiblir de savoir un peu de mathématiques, de biologie, d'histoire, etc., ***et d'en faire usage*** lorsque la situation semble l'exiger ?

➔ Une enquête est presque à tout coup ***codisciplinaire*** : non seulement elle met à contribution ***plusieurs*** disciplines, mais celles-ci y interviennent ***en synergie*** – elles y « travaillent » ***de concert*** (ce que marque le préfixe *co*-). Mais cette codisciplinarité est source de difficultés dans une société où la valeur épistémologique la mieux affirmée, la plus prégnante est celle de la (prétendue) « spécialité » : un tel est biologiste, tel autre ne l'est pas ; il est pharmacologue, ou historien de la Grèce ancienne, ou probabiliste, ou angliciste, etc. Ou peut-être rien de tout cela ! Ce que suppose le développement et la diffusion d'une culture de l'enquête codisciplinaire, c'est donc une véritable ***ré-appropriation praxéologique***, une reconquête du droit à la connaissance et à son usage ***légitime mais responsable*** – ce qui suppose notamment d'adopter une certaine ***éthique de la mise en question*** des réponses auxquelles on parvient.

d) La possibilité pratique de conduire des enquêtes se concrétisant en des PER de plus ou moins longue durée a été considérablement accrue par l'avènement et le développement impétueux de l'***Internet***, dont on sait qu'il a pour beaucoup de gens encore le statut de « mauvais objet », mais qui a notamment le mérite de dénoncer *de facto* quelques-uns des principaux obstacles qui se sont installés dans nos cultures – souvent par le truchement de l'éducation scolaire – en matière de ***rapport à la connaissance***.

➔ Dans ce qui suit, c'est avec en tête ce recours à l'Internet dans un travail d'enquête codisciplinaire que l'on examinera la présentation des principaux obstacles d'origine scolaire, et cela à travers les praxéologies didactiques – les « dialectiques », dont quelques-unes ont déjà été mentionnées – qui doivent permettre de les surmonter.

1) La première dialectique est celle ***du sujet et du hors sujet*** : contre le postulat scolaire du plus court chemin, qui ne conduit qu'à un but connu et déterminé ***à l'avance***, elle pousse, dans une recherche en principe ***ouverte***, à risquer le hors sujet tant en matière de recherche documentaire par exemple que dans le choix des questions Q_1, Q_2, etc., engendrées par l'étude de Q, et dont on décidera ou non d'entamer ou de poursuivre l'étude.

2) La deuxième dialectique est celle ***du parachutiste et du truffier*** : contre le double habitus scolaire de la rareté documentaire et de la recherche de l'adéquation immédiate du document au projet d'étude et de recherche, elle conduit à « ratisser » de ***vastes zones***, où l'on sait *a priori* qu'on ne trouvera pas grand-chose, mais où pourra advenir ***de l'inattendu***, et où l'on apprendra à repérer les rares « pépites », souvent peu visibles, qui feront progresser la recherche.

3) La troisième dialectique est celle ***des boîtes noires et des boîtes claires*** : contre le primat donné à la connaissance ***déjà*** disponible, elle invite à donner le primat à la connaissance ***pertinente***, quel que soit *a priori* son statut au regard des savoirs enseignés, à limiter au nécessaire la clarification (les boîtes réputées « claires » sont ***toujours*** des boîtes ***grises***), à prendre donc le risque, ponctuellement,

– de clarifier des boîtes noires situées à la limite du curriculum officiel,

– de laisser dans l'obscurité ce que, dans le curriculum familier, l'on vise par ailleurs à clarifier,

– de traquer les boîtes « invisibles » parce que « transparentes », pour ***déconstruire les évidences*** de la culture de l'institution ***chaque fois que c'est utile***.

4) La quatrième dialectique est celle qu'on peut appeler, classiquement, ***de la conjecture et de la preuve***, mais qu'on nommera, dans une perspective plus large, ***des médias et des milieux*** : contre la mise à l'épreuve plus ou moins réglée à l'avance d'assertions réputées sûres en vertu surtout de l'autorité de l'institution, elle engage à soumettre les assertions obtenues à la critique des diverses dialectiques et à évaluer le ***degré d'incertitude*** d'une assertion donnée.

5) La cinquième dialectique est celle ***de la lecture (= de « l'excription ») et de l'écriture (= de l'inscription)*** : contre le recopiage formel de textes où ont été inscrites des réponses $R^{\lozenge}$ que la mise en texte a « dévitalisées », elle convie à entrer dans la dialectique de la lecture « excriptrice », qui redonne vie aux réponses $R^{\lozenge}$ déposées dans les documents disponibles, et de l'écriture « inscriptrice » d'une réponse propre $R^{\heartsuit}$ qui prend forme peu à peu ***par le croisement de plusieurs niveaux d'écrit*** (carnet de bord, notes de synthèse, glossaire, production finale).

6) La sixième dialectique est celle ***de la diffusion et de la réception*** : contre la tentation de ne pas défendre sa réponse *R*, supposée par avance connue et reconnue par l'institution où elle est produite, contre l'opportunisme à l'endroit de *R* afin de complaire à qui l'on s'adresse, elle invite à ***défendre R*** sans infidélité au travail accompli, mais dans l'attention à ce qu'autrui en peut recevoir.

➔ Ces notations ont été rédigées, à l'origine, comme un guide « praxéologique » à l'adresse de professeurs ayant à encadrer des TPE, des ***travaux personnels encadrés***, en classe de première. Sur la question de ce qu'ont été les TPE, de ce qu'ils sont, de ce qu'ils pourraient être, de ce qu'ils devraient être de tel point de vue, on laissera le lecteur mener son enquête *ab ovo usque ad mala*, « depuis les œufs jusqu'aux pommes », c'est-à-dire d'un bout à l'autre (du repas), selon l'expression d'Horace dans les *Satires* (I, 3).

Annexe

« Internet et moi » : une ébauche d'analyse des données recueillies

1. Rappelons d'abord les questions sur lesquelles les enquêtés devaient s'interroger (de deux façons, successivement) :

Question 1. *Internet, qu'est-ce que c'est ?*

Question 2. *Comment les ordinateurs sont-ils connectés à Internet ?*

Question 3. *Qu'est-ce qu'un site Internet ?*

Question 4. *Comment les informations circulent-elles sur Internet ?*

Question 5. *Peut-on pirater Internet ?*

Question 6. *Comment fonctionne la messagerie électronique ?*

Dans ce qui suit, ces questions sont désignées respectivement par les notations Q_1, Q_2, Q_3, Q_4, Q_5, Q_6.

2. Les enquêtés étaient au nombre de 11 ; ils sont désignés dans ce qui suit par les notations J_1, …, J_{11} (la lettre *J* est l'initiale du mot « juge »). Ces onze juges avaient d'abord à estimer *a priori*, sur une échelle de 0 à 4, et pour chacune des six questions – censées leur être posées par un enfant d'une dizaine d'années – la qualité de leur réponse éventuelle au petit curieux. Leurs choix sont présentés dans le tableau ci-après, augmentés : 1) pour chaque question, du total des scores recueillis par la question et du pourcentage du total maximal (égal à $4 \times 11 = 44$) ainsi obtenu ; 2) pour chaque juge, du total des scores attribuées aux questions et du pourcentage du total maximal (égal à $4 \times 6 = 24$) ainsi obtenu.

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Total	%
J_1	3	2	3	3	2	3	16	66,7
J_2	2	3	1	1	3	2	12	50,0
J_3	3	2	2	2	4	2	15	62,5
J_4	3	3	3	2	1	2	14	58,3
J_5	3	2	1	3	2	0	11	45,8
J_6	2	1	3	1	0	4	11	45,8
J_7	2	0	3	2	0	2	9	37,5
J_8	2	3	2	1	1	1	10	41,7
J_9	3	1	3	1	3	3	14	58,3
J_{10}	2	1	2	1	0	1	7	29,2
J_{11}	3	2	4	2	4	3	18	75,0
Total	28	20	27	19	20	23	137	
%	63,6	45,5	61,4	43,2	45,5	52,3		

➔ On considère ici que, plus le score assigné à une question est élevé, plus la ***facilité*** estimée de (répondre à) la question est élevée : on parlera de P-facilité (*a priori*), où facilité estimée (*a priori*) de la performance (d'où le préfixe P). Les données recueillies conduisent à classer les questions du point de vue de la P-facilité de la façon suivante :

$$Q_4\ (43{,}2) \prec Q_2 \sim Q_5\ (45{,}5) \prec Q_6\ (52{,}3) \prec Q_3\ (61{,}4) \prec Q_1\ (63{,}6).$$

Les deux questions jugées les plus P-faciles sont Q_1 (*Internet, qu'est-ce que c'est ?*) et Q_3 (*Qu'est-ce qu'un site Internet ?*). Vient ensuite Q_6 (*Comment fonctionne la messagerie électronique ?*). Ces trois questions s'opposent au bloc des questions Q_2, Q_4, Q_5 qui « n'ont pas la moyenne » (43,2 et 45,5).

➔ Une interprétation simple du constat précédent est possible : les questions Q_1, Q_3, Q_6 ont trait à des réalités proches de la pratique des juges : on ne peut guère « pratiquer » un tant soit peu l'Internet sans penser que l'on sait en gros ce que c'est qu'Internet (Q_1), ce qu'est un site Internet (Q_3) et comment fonctionne la messagerie électronique (Q_6). En revanche, la manière dont les ordinateurs sont connectés à Internet (Q_2), la façon dont les informations y circulent (Q_4) ou les possibilités de piratage (Q_5) sont hors du champ d'activité du groupe des juges.

3. Les onze juges ont eu ensuite à estimer *a posteriori* (de façon imaginaire), sur une échelle de 0 à 4, et pour chacune des six questions leur manque de connaissances à propos de ces questions. Le tableau ci-après a la même structure que le tableau précédent.

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Total	%
J_1	1	1	1	2	3	1	9	37,5
J_2	2	1	3	2	1	3	12	50,0
J_3	1	2	2	3	1	4	13	54,2
J_4	1	1	1	2	3	2	10	41,7
J_5	1	3	3	2	2	4	15	62,5
J_6	2	3	1	3	3	0	12	50,0
J_7	2	4	2	3	4	2	17	70,8
J_8	3	3	4	3	3	3	19	79,2
J_9	1	3	1	2	2	1	10	41,7
J_{10}	2	3	3	4	4	3	19	79,2
J_{11}	1	2	0	3	2	1	9	37,5
Total	17	26	21	29	28	24	145	
%	38,6	59,1	47,7	65,9	63,6	54,5		

➔ On regarde ici le score assigné comme estimant (*a posteriori*) la C-exigence de la question considérée, c'est-à-dire le caractère exigeant (estimée *a posteriori*) de telle question en termes de compétence (d'où le préfixe C).

➔ Les données recueillies conduisent à classer les questions du point de vue de leur C-exigence de la façon suivante :

$$Q_1\ (38{,}6) \prec Q_3\ (47{,}7) \prec Q_6\ (54{,}5) \prec Q_2\ (59{,}1) \prec Q_5\ (63{,}6) \prec Q_4\ (65{,}9).$$

La première observation est évidemment que – de façon très logique – ce classement selon la C-exigence inverse le classement selon la P-difficulté (à ce détail près qu'il distingue les questions Q_2 et Q_5, qui ne l'étaient pas précédemment).

➔ Une seconde notation porte sur ce fait que, hormis peut-être s'agissant de Q_1, les « manques » estimés semblent quantitativement importants : ils s'étagent, grosso modo, entre 48 % et 66 %. La situation observée porte à penser que « l'équipement praxéologique » des juges en matière d'Internet est regardé par eux comme fragile, même si, mis devant l'obligation d'agir (en répondant aux questions d'un enfant), ils estiment ne pas devoir faire si mal.