Organisation et gestion de données

Jean-Claude DUPERRET - IREM de Reims

Cette partie du programme est "naturellement" celle où les mathématiques et la réalité vont être le plus en interaction. Si les mathématiques ne sont pas la réalité, elles permettent d'expliquer, de valider, de modéliser des situations issues du monde "réel", et, inversement, ces situations permettent de motiver l'apprentissage des mathématiques.

C'est dans cette partie que l'élève va de façon privilégiée développer des aptitudes à trier, ranger, transformer des informations, en s'appuyant sur de fréquents changements de registres. Partant d'un texte, souvent écrit en français, comportant un certain nombre de données chiffrées, il devra déjà organiser ces données, par exemple sous la forme de tableau ou de graphique, deux cadres dont il faut développer l'interactivité dès la sixième. Puis des calculs permettront de transformer ces données, de les synthétiser. De même, les résultats pourront eux aussi être donnés dans différents registres suivant la nature du problème étudié : texte, tableau, graphique, résultat numérique...

C'est enfin dans cette partie qu'apparaît le mieux le rôle social et culturel des mathématiques. La lecture, l'interprétation, l'utilisation de diagrammes, tableaux, graphiques, leur analyse critique aident l'élève à mieux comprendre le monde dans lequel il vit, et, en cela, contribuent à son éducation civique. La liaison avec l'enseignement d'autres disciplines, en particulier sciences de la vie et de la terre, géographie, technologie lui permettent de mieux intégrer ses connaissances dans une vision plus large.

Comment les mathématiques vont-elles participer à la réalisation d'un tel enjeu ? En proposant des modèles de traitement et de validation ; et, parmi tous ces modèles, celui qui va permettre de gérer le maximum de situations est la proportionnalité. Si ce modèle est un des plus développés au collège, il reste cependant un modèle difficile et à haut risque. Il faut à son sujet rappeler que son enseignement relève de beaucoup de disciplines, et que les mathématiques n'en sont que le lieu de systématisation. Il faut aussi éviter le risque du "tout proportionnel" et du "tout linéaire" en proposant des situations de non proportionnalité.

Pour les élèves le concept de proportionnalité est difficile. Il est donc normal que tout enseignant, dès la sixième, se pose la question : pourquoi ?

Une première raison est peut-être la tendance de notre enseignement à vouloir trop rapidement tout unifier. Illustrons cela par deux exemples utilisant les pourcentages : dans une classe, il y a 8% d'élèves redoublants ; les cigarettes ont augmenté de 8% . Si ces deux situations relèvent de la proportionnalité, si les deux utilisent le même pourcentage, ce pourcentage n'a pas du tout la même signification par rapport à la situation. Peut-on en effet considérer avec le même regard un pourcentage exprimant le rapport d'une partie à un tout et un pourcentage exprimant une augmentation ? Il faut donc accepter que dans la tête des élèves, il y ait des proportionnalités correspondant à des champs de problèmes différents. Le rôle de l'enseignement des mathématiques est à terme (quatrième-troisième) de les unifier dans un grand modèle.

Partant de ce constat, quelles réflexions et quelles activités proposer aux élèves ? Il est bien entendu impossible d'être exhaustif, et les trois thèmes proposés ci-après ne se veulent pas des références, mais des exemples. Le but y est de développer chez l'élève cet esprit critique d'analyse par rapport à des situation "réelles", tout en développant des modèles de travail, en particulier la proportionnalité.

THEME 1 : Un pourcentage ne peut donner que ce qu'il a !

Un certain nombre de "mots" peuvent permettre de "représenter" une partie par rapport à un tout : proportion, rapport, fraction, fréquence, pourcentage. Chacun d'eux a sa spécificité, que nous n'étudierons pas ici. Nous choisissons donc le plus neutre : le pourcentage, en considérant que les règles d'utilisation et d'action qui s'y rattachent ont été travaillées.

Réflexion à mener avec les élèves :

Je veux comparer les collèges Albert Camus et Paul Langevin au niveau des redoublements en classe de troisième en 1996. A Albert Camus, il y a eu 15 redoublants, à Paul Langevin 12. Une conclusion immédiate est qu'il y a plus de redoublants à Camus qu'à Langevin. Mais je regarde les effectifs de chacun des collèges : il y avait 125 élèves en troisième à Camus, et 80 à Langevin. Un calcul (à mener avec les élèves) donne donc un taux de doublement de 12% à Camus et de 15% à Langevin. Le taux de doublement est donc plus élevé à Langevin qu'à Camus.

L'exemple précédent montre que pour comparer deux populations, on peut le faire soit de façon absolue, soit de façon relative. On peut parfois trouver dans le texte de la question si on attend plutôt l'une ou l'autre des comparaisons. Dans l'exemple précédent, si on demande : dans quel collège y a-t-il le plus de "redoublants", on attend une comparaison absolue ; mais si on s'intéresse au "redoublement", on attend plutôt la comparaison relative (même si le mot "taux" n'est pas prononcé). Cela n'est pas toujours très clair : par exemple, si je veux savoir de deux pays lequel est le plus peuplé, est-ce que je m'intéresse à la densité de population, ou au nombre d'habitants !

Activité :

Dans une petite ville de France, il y a deux collèges : le collège Pierre Brossolette et le collège Gaston Bachelard. Le premier a 45% de garçons et 55% de filles. Le second a 60% de garçons et 40% de filles. Y a-t-il plus de garçons ou de filles scolarisés en collège dans cette ville ?

Malgré la réflexion précédemment menée, il se trouvera certainement des élèves pour répondre spontanément : les garçons ! D'autres au contraire seront plus prudents. Certains iront peut-être jusqu'à dire qu'il faut connaître les effectifs de chaque collège pour répondre ! On pourra de toute façon relancer (et boucler !) cette activité en proposant trois types d'effectifs. Par exemple :

* 420 élèves à Brossolette et 360 à Bachelard, ce qui donnera 405 garçons et 375 filles.
* 520 élèves à Brossolette et 260 à Bachelard, ce qui donnera 390 garçons et 390 filles.
* 740 élèves à Brossolette et 300 à Bachelard, ce qui donnera 513 garçons et 527 filles.

THEME 2 : Faut-il joindre les points ?

Les graphiques sont devenus des outils fondamentaux de communication. Il est hors de question en quelques lignes d'entrer véritablement dans les problèmes soulevés par leur enseignement. Pour ceux qui souhaiteraient aller plus loin dans cette réflexion, ils trouveront matière dans la Brochure Inter Irem 1er cycle "Des chiffres et des lettres" , en particulier avec le travail de l'IREM de Rouen.

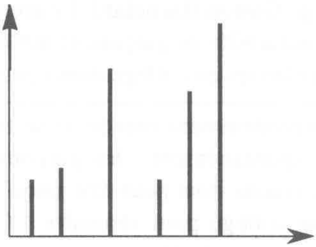
Réflexion à mener avec les élèves :

On propose aux élèves un certain nombre de graphiques , bons ou moins bons, clairs ou surchargés, "vrais" ou "faux". A partir de cet échantillon, on peut dégager trois qualités nécessaires à une bonne réalisation.

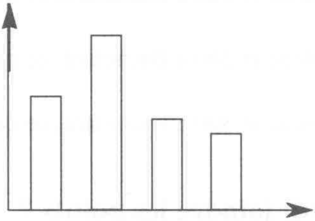
* *la lisibilité* : il ne faut pas que, par une surcharge d'informations, l'essentiel de ce qui doit être communiqué n'apparaisse pas ;

* *l'autosuffisance* : il faut qu'il puisse être lu, interprété et compris sans avoir recours à d'autres sources d'information ;
* *la validité* : il faut qu'il puisse être validé (ou invalidé) par des règles mathématiques (ex : proportionnalité des effectifs et des longueurs, des effectifs et des aires, ...).

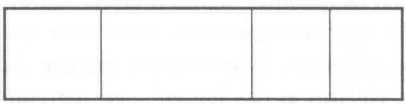
Pour une même situation, il peut être intéressant de produire plusieurs graphiques, et de comparer les différences de qualité d'information qu'ils donnent. Par exemple, si on étudie la série des notes obtenues à un devoir, on peut, dans un premier temps, faire un diagramme en bâtons :



Puis, en regroupant les notes suivant quatre catégories ; de 0 à 5, de 6 à 10, de 11 à 15, de 16 à 20, on peut faire un diagramme en barre :



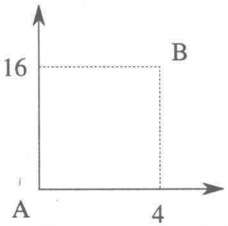
et un diagramme en bande représentant ces quatre catégories d'élèves :



Le diagramme en bâtons est le plus précis ; le diagramme en barre donne une photographie rapide de la classe ; le diagramme en bande fait apparaître la place de chaque catégorie par rapport à la classe. Ces graphiques ci-dessus ne représentent évidemment pas la même situation.
Si on peut passer du diagramme en bâtons au diagramme en barre, le retour est impossible, car les informations sont insuffisantes. C'est là un problème important : quelles informations puis-je m'autoriser à ajouter en fonction de celles que j'ai et de la connaissance de la situation ?

Activité :

Voici quatre situations :



1) Le prix d'un croissant est 4 F.
2) Le prix d'un mètre de tissu est 4 F.
3) Quel est le périmètre d'un carré en fonction de la longueur du côté (en cm) ?
4) Quelle est l'aire d'un carré en fonction de la longueur du côté (en cm) ?

Question 1 : En graduant convenablement dans chaque cas les axes, vérifier que les points A et B appartiennent à la représentation graphique de chacune des situations.
Question 2 : Pour chaque situation, est-il "correct" de joindre A et B à la règle ?

Tracer le segment [AB], c'est considérer que tous les points de [AB] appartiennent à la représentation graphique de la situation. Un débat avec les élèves permet de voir que :

- *Dans la situation 1* : seuls les points à coordonnées entières "ont du sens", donc tracer [AB] n'est pas "correct" (bien que l'on soit dans une situation de proportionnalité !).
- *Dans la situation 2* : le problème ici est bien plus délicat. Reste-t'on sur l'aspect "concret" ? Si 1,60 m de tissu acheté a du sens, 1,63 m est déjà d'une précision guère "réelle", et que dire de 1,63297 m, ou encore $\sqrt{2}$ m et Π m ! Idéalise-t on ? Alors tout ce qui précède est "correct", et tous les points de [AB] conviennent.
  Il faut aussi , au delà du statut des nombres employés tenir compte de la précision de la lecture possible du graphique.
- *Dans la situation 3* : cette situation est déjà ici "idéalisée", et tous les points de [AB] conviennent.
- *Dans la situation 4* : quelques contre-exemples permettront de vérifier que les points ne sont pas alignés (on n'est pas dans une situation de proportionnalité).

Theme 3 : " Au hasard ! "

Les médias nous "baignent" quotidiennement dans des sondages, des simulations, des enquêtes statistiques... Tout cela semble passer par le mot magique "échantillon". Pourquoi travailler sur des échantillons ? Comment les constituer ? Quelle crédibilité accorder aux résultats qu'on en tire ?

Réflexion à mener avec les élèves

Je suis un fabricant de "pétards". Voyant arrivé le 14 juillet, je m'inquiète de savoir si mes pétards sont en bon état. Pour être sûr de l'état de mon stock, je peux évidemment tous les essayer. Je pourrais alors affirmer, j'en avais 97% de bons. Mais je n'aurai plus de stock !

Je suis un homme politique, et j'aimerais "prévoir" les résultats des prochaines élections nationales. Je peux évidemment interroger tous les français votants. Mais comment les contacter tous ? Où ? Combien de temps me faudra-t-il ?

La réponse est de travailler sur une partie de la "population". Dans le premier cas, on essayera par exemple un pétard sur 150. Dans le second cas, on se contentera d'interroger 1000 personnes. Une telle partie est un échantillon. Une première question fondamentale est : " Qu'est ce qu'un bon échantillon ? ". C'est une partie de la population qui, pour la question que j'étudie, me donne sensiblement les mêmes réponses que la population toute entière.

Dans le premier exemple, il faudra donc que l'échantillon me donne 97% de bons pétards. Dans le second exemple, il devra me donner la photographie du résultat des élections.

Une question corollaire encore plus fondamentale est : "Comment faire un bon échantillon ? ". La réponse immédiate est : "par un tirage au hasard ? ". Mais comment faire un tirage au hasard ? Par exemple, pour le cas des élections, je peux décider de me mettre sur une place dans le centre d'une ville, et d'interroger les 1000 premiers passants. On voit bien tout de suite que les résultats ne seront pas les mêmes si je suis dans une ville de province, dans une grande ville, ou dans une ville de banlieue.

En fait, pour éviter les "catastrophes" que pourrait amener le hasard sur l'enquête menée, on décide en général de constituer un échantillon à partir de critères objectivement définis. Une méthode particulièrement employée est la "méthode des quotas", qu'on trouve développée par l'IREM de Montpellier dans la brochure "*Des chiffres et des lettres*" déjà citée. Illustrons cette méthode sur un exemple.

Activité :

> Je voudrais connaître les émissions télévisées préférées du collège Georges BRASSENS. Mais ce collège a 800 élèves, et, pour des raisons de temps, je ne peux interroger individuellement ces 800 élèves. Je décide donc de n'interroger qu'un échantillon de 60 élèves. Comment constituer un échantillon représentatif ?

Après la réflexion précédemment menée, on peut rapidement convenir avec les élèves qu'un échantillon de soixante élèves tirés au hasard ne serait peut-être pas représentatif. Il faut donc dégager des critères objectifs de constitution de cet échantillon. Deux peuvent être assez facilement mis en avant : le sexe et l'âge. Si une fille de dix ans a une admiration sans borne pour "Dorothée", la préférence d'une adolescente de quinze ans ira plutôt à "Hélène et les garçons", tandis qu'un garçon de quinze ans sera peut-être plus sensible à la retransmission de matchs de football.

Voici donc la répartition des élèves du collège en fonction de ces deux critères :

	Filles	**Garçons**
1985	50	40
1984	100	120
1983	90	110
1982	80	70
1981	60	80

On peut alors traduire cette répartition en pourcentages :

	Filles	**Garçons**
1985	6,25%	5%
1984	12,5%	15%
1983	11,25%	13,75%
1982	10%	8,75%
1981	7,5%	10%

On applique alors ces pourcentages à soixante, en gérant les problèmes d'arrondi ! (En particulier, on essaie de garder l'équilibre de la population du collège par corrections successives des arrondis.)

	Filles	**Garçons**
1985	4	3
1984	7	9
1983	7	8
1982	6	5
1981	5	6

On va donc constituer un échantillon en prenant (et là vraiment au hasard !) 4 filles de 1985, 3 garçons de 1985, ..., 6 garçons de 1981.

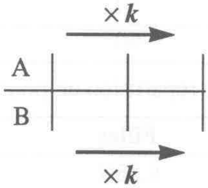
La gestion d'une telle activité pose de façon très forte le problème des arrondis, et ce n'est pas le moindre des débats à avoir avec les élèves. Il faut aussi noter dans l'exemple proposé, que les nombres choisis ne sont pas le fruit du hasard ! Il vaut évidemment mieux travailler avec les données réelles de son collège.

******

Les trois thèmes que nous venons de traiter utilisent essentiellement la proportionnalité comme méthode de travail (en mettant évidemment à part la relation liant l'aire d'un carré à son côté).

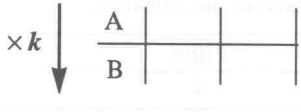
Cette forte présence nous invite à revenir sur ce modèle, et, au delà des raisons pédagogiques, d'analyser en quoi il est difficile d'un point de vue purement mathématique. Je m'appuierai pour cela sur une réflexion que m'a proposée Rudolf BKOUCHE (IREM de Lille) : on peut donner deux définitions de la proportionnalité dont l'équivalence ne va pas de soi.

La première s'appuie sur les *rapports internes* :



Deux espèces de grandeurs A et B (possédant la propriété d'additivité) étant données, une relation de proportionnalité entre ces deux grandeurs est une correspondance qui conserve les "rapports internes". Par exemple, dans un mouvement uniforme, le rapport des distances parcourues est égal au rapport des durées du parcours.

La seconde s'appuie sur les *rapports externes* :



Les grandeurs A et B étant mesurées, le rapport de la mesure d'une grandeur de l'espèce B par la mesure de la grandeur correspondante de l'espèce A est constant. Par exemple, dans un mouvement uniforme, le rapport de toute distance parcourue à la durée correspondante est constant. On retrouve ici la notion de coefficient de proportionnalité, qui, dans l'exemple proposé, s'appelle la vitesse.

La première définition pose un problème fondamental : comment définir un rapport interne ? La réponse est *la mesure,* qui n'est rien d'autre qu'une correspondance proportionnelle entre grandeur et nombres. De façon plus précise la mesure d'une grandeur est le nombre qui exprime le rapport de cette grandeur à la grandeur unité.

Les hésitations des élèves sur le lien entre ces deux définitions nous rappellent que leur équivalence est loin d'être évidente, et c'est le rôle de l'enseignement des mathématiques que de l'établir. La relation entre rapports externes et rapports internes se traduit en effet par l'équivalence des relations $\frac{a}{b} = \frac{c}{d}$ et $\frac{a}{c} = \frac{b}{d}$. Si une façon commode et usitée de régler ce problème est le passage par les produits en croix, il faut bien voir que ceci consiste à remplacer un concept par une procédure.

Et nous touchons là un point sensible : très rapidement pour les élèves, la proportionnalité n'apparaît que comme une suite de procédures. Si on ne peut nier l'intérêt de développer des mécanismes qui permettent une économie dans le travail, il ne faut cependant pas que ces mécanismes deviennent l'objectif principal d'apprentissage.

Nous revenons là sur la réflexion du départ : si les mathématiques donnent des outils pour résoudre des problèmes, elles ont aussi un rôle d'analyse critique. Il faut donc proposer des situations où l'élève ne doit pas se contenter d'**agir,** mais doit aussi **réfléchir.**