

L'organisation et la gestion de données au cycle 4 : quelles difficultés ?

The organization and management of data in cycle 4: what are the challenges?

Anne Crumière¹

UMR EFTS, Université Toulouse Jean Jaurès (ESPE), France

<https://orcid.org/0000-0001-6575-1364>

Gisèle Cirade²

UMR EFTS, Université Toulouse Jean Jaurès (ESPE), France

<https://orcid.org/0000-0002-1894-817X>

Résumé

Nous nous appuyons sur des observations réalisées dans le cadre de la formation initiale des professeurs de mathématiques pour mettre au jour certains problèmes de la profession, en nous centrant sur le cas de l'organisation et de la gestion de données au cycle 4 (élèves de 12-15 ans). Nous mettons en évidence une péjoration du savoir lié à ce secteur, qui conduit à une instabilité de certaines notions et à de nombreuses imprécisions ou inadéquations. Nous dégageons aussi les difficultés qu'ont les professeurs à proposer des situations riches à étudier et à donner à la classe la possibilité de prendre des chemins variés, ce qui va de pair avec l'absence de gestes propres à l'enquête dans leurs praxéologies de direction d'étude.

Mots-clés: Enquête, Praxéologies de direction d'étude.

Abstract

We draw on observations made in the frame of the initial training of mathematics teachers, aiming to highlight some problems of the profession, with focus on the case of data organisation and data management in cycle 4 in France (12-15-year-old students). We demonstrate a deterioration of the knowledge linked to this sector, that results in instabilities of certain notions as well as many inaccuracies or inadequacies. We also

¹ anne.crumiere@univ-tlse2.fr

² gisele.cirade@univ-tlse2.fr

identify the difficulties encountered by professors in proposing rich enough situations to study and in giving the class the opportunity to take a variety of paths, which are in tandem with the absence of actions specific to the investigation in their praxeologies of study direction.

Keywords: Investigation, Praxeologies of study direction.

L'organisation et la gestion de données au cycle 4 : quelles difficultés ?

Le secteur « Organisation et gestion de données » a été introduit dans les programmes du collège dès 1996 (Ministère de l'Éducation nationale [MEN], 2004). Reconduit à chaque nouveau programme depuis cette date, il prend de l'ampleur – notamment avec l'introduction de l'échantillonnage, de l'estimation et des intervalles de fluctuation au lycée – tout en ayant pourtant beaucoup de mal à exister en tant que tel. Pour étudier ce que nous considérons comme un *problème de la profession* (Chevallard, 2013), nous commençons par dégager certaines difficultés du métier relativement à l'enseignement de ce secteur au cycle 4. Pour ce faire, sachant que les difficultés auxquelles se heurtent les débutants constituent presque toujours des *révélateurs* de celles qui sont rencontrées par la profession (Cirade, 2006), nous choisissons de nous appuyer sur des observations réalisées dans le cadre de la formation initiale des professeurs de mathématiques – on notera d'ailleurs que notre analyse nous amène à examiner des manuels scolaires et des progressions communes dans les établissements, et que cela met en évidence, s'il en est besoin, que les difficultés repérées ne sont pas l'apanage des débutants.

Toutes les observations mentionnées dans cet article ont été réalisées dans le cadre de la formation des élèves professeurs de deuxième année à l'ESPE Toulouse Midi-Pyrénées. Elles s'appuient sur des activités proposées dans les deux premiers mois de la formation en didactique, à une période où les objectifs essentiels sont, d'une part, l'analyse praxéologique des mathématiques constituant l'enjeu de l'étude – dans les programmes, les manuels, les comptes rendus de séance, etc. – et, d'autre part, la conception d'une activité d'étude et de recherche.

Choix d'un support d'activité d'étude et de recherche

Commençons par examiner ce que dit le programme du cycle 4 à ce sujet. L'introduction du domaine « Organisation et gestions de données, fonctions » indique

que, « au cycle 4, les élèves apprennent à utiliser une représentation adaptée de données pour en faire une interprétation critique » (MEN, 2015, p. 372). Nous nous arrêterons sur le premier des quatre attendus de fin de cycle, « Interpréter, représenter et traiter des données » (voir figure 1) :

Figure 1

Extrait du programme du cycle 4. (p. 373)

Connaissances et compétences associées	Exemples de situations, d'activités et de ressources pour l'élève
Interpréter, représenter et traiter des données	
Recueillir des données, les organiser. Lire des données sous forme de données brutes, de tableau, de graphique. Calculer des effectifs, des fréquences. » Tableaux, représentations graphiques (diagrammes en bâtons, diagrammes circulaires, histogrammes). Calculer et interpréter des caractéristiques de position ou de dispersion d'une série statistique. » Indicateurs : moyenne, médiane, étendue.	Utiliser un tableur, un grapheur pour calculer des indicateurs et représenter graphiquement les données. Porter un regard critique sur des informations chiffrées, recueillies, par exemple, dans des articles de journaux ou sur des sites web. Organiser et traiter des résultats issus de mesures ou de calculs (par exemple, des données mises sur l'environnement numérique de travail par les élèves dans d'autres disciplines) ; questionner la pertinence de la façon dont les données sont collectées. Lire, interpréter ou construire un diagramme dans un contexte économique, social ou politique : résultats d'élections, données de veille sanitaire (par exemple consultations, hospitalisations, mortalité pour la grippe), données financières relatives aux ménages (par exemple impôts, salaires et revenus), données issues de l'étude d'un jeu, d'une œuvre d'art...

On peut dégager divers types de tâches relatifs à cet attendu de fin de cycle en se référant (schématiquement) aux différentes phases de l'étude d'une situation du monde (Chevallard, 1989, 2007) selon que le type de tâches relève de la modélisation (recueillir et organiser des données), du travail du modèle (lire des données, calculer des effectifs et des fréquences, calculer des caractéristiques de position ou de dispersion) ou du retour à la situation (interpréter des caractéristiques de position ou de dispersion). On notera que le recueil, l'organisation et la lecture des données constituent une reprise d'étude du cycle 3.

La recherche d'un support d'activité

Après un travail autour de l'organisation mathématique de ce secteur, il a été proposé aux élèves professeurs un atelier sur l'organisation didactique, où les équipes avaient à considérer la consigne suivante³ : « Choisir un énoncé pour démarrer une 'activité' permettant de faire émerger dans la classe l'organisation mathématique ponctuelle autour du type de tâches 'Déterminer une médiane'. » Commençons par examiner les trois propositions que nous avons reproduites en annexe (équipes A, B et C), qui s'appuient sur des situations du monde extramathématique.

L'équipe A propose de recueillir les pointures des 29 élèves de la classe. Les étapes se succèdent : on recueille les données, on range les valeurs dans l'ordre croissant, on repère la note du 15^e élève dans le tableau ordonné et on lit la pointure associée. Mais, à aucun moment, la raison de la détermination de la pointure du 15^e élève n'est évoquée. Par ailleurs, on peut lire en *nota bene* : « si classe d'effectif pair : au départ, le professeur ajoute sa pointure ». Cette volonté de travailler, au moins dans un premier temps, sur un effectif impair, sera reprise par quatorze des quinze équipes d'élèves professeurs. Elle nous semble révélatrice d'un *évitement* du cas où l'effectif est pair dans l'exploration du type de tâches et nous faisons l'hypothèse que, quelquefois, pour le cas pair, la technique et la technologie seront simplement fournies par le professeur – ce qui limite alors très fortement le *topos* de l'élève dans la réalisation du moment exploratoire et du moment technologico-théorique. L'équipe B considère le poids de 15 judokas : le but de l'activité proposée est de répartir les compétiteurs en deux poules, les « légers » et les « lourds », de telle sorte que la répartition soit la plus équitable possible et, s'il y a bien une question problématique, elle ne correspond pas du tout aux pratiques usuelles dans les

³ Une formulation légèrement différente a été proposée à un autre groupe, mais les différences sont inessentiels pour ce que nous regardons ici

compétitions de judo. L'équipe C, quant à elle, considère la taille de 11 vidéos et propose de déterminer la médiane, mais cela ne vient répondre à aucune question.

Dans chacun de ces cas, aucune question problématique digne de ce nom n'est posée – la médiane constitue ici un *enjeu imposé* par le professeur et non par l'étude d'une question –, les énoncés sont guidés par un découpage en question très fermées, les données sont peu nombreuses et les traitements afférents peuvent être réalisés à la main. Concernant le choix du support d'activité, notons tout d'abord la propension des élèves professeurs à créer leur propre support, ainsi que le signale l'équipe A dans sa production : « Créé de toutes pièces ». Mais un autre phénomène doit aussi être signalé, qui est très prégnant : peu d'équipes suivent les recommandations du programme et des ressources thématiques enjoignant les professeurs à choisir des situations « authentiques », ainsi que le mentionne la ressource thématique intitulée « Interpréter, représenter et traiter des données », qui précise notamment que « cette partie du programme s'appuie largement sur l'étude de situations authentiques issues de la vie courante, de mesures obtenues dans d'autres disciplines par les élèves eux-mêmes, de données scientifiques ou économiques disponibles sur Internet » (MEN, 2016, p. 6). Sur quinze équipes, seules deux envisagent de se situer dans un tel contexte, en proposant de travailler sur les salaires, mais la situation porte sur une petite entreprise, avec 9 employés dans un cas et 12 dans l'autre.

Figure 2 –

Extrait de la ressource thématique susmentionnée. (p. 5)

Exemple

Source : [Salaires dans le secteur privé et les entreprises publiques](#)

Le salaire mensuel net moyen en France en 2013, dans le privé ou en entreprise publique, s'est établi à 2202 euros. Le salaire mensuel net médian pour la même année s'élevait à 1772 euros.

La ressource thématique, quant à elle, renvoie à des données bien plus amples (Figure 2).

Le travail sur un support d'activité

Dans un autre atelier, il a été proposé aux élèves professeurs de mener eux-mêmes une étude statistique à partir de la consigne suivante⁴ :

Quand on dit qu'une commune française n'a pas beaucoup d'habitants, ça veut dire qu'elle en a combien ? On considérera le fichier Excel disponible sur la page suivante du site de l'INSEE : <https://www.insee.fr/fr/statistiques/2864136> et on utilisera les connaissances statistiques au programme du collège. Le fichier Excel contient des données sur la population des communes françaises.

Au cycle 4, on peut envisager une technique s'appuyant sur le calcul des fréquences cumulées⁵. Un premier geste consiste à classer les communes dans l'ordre croissant ou décroissant de leur population ; compte tenu de la nature de la question, on choisit le tri croissant. Les premières communes qui apparaissent n'ont pas beaucoup d'habitants, mais le problème se pose de savoir où « arrêter le curseur ». On peut alors dénombrer les communes qui ont moins de N habitants, en prenant quelques valeurs de N , mais on voit assez vite que travailler avec les effectifs n'est pas très probant : on a par exemple 3391 communes qui ont moins de 100 habitants, mais est-ce peu ou beaucoup ? C'est par rapport au nombre de communes de la France qu'il va falloir se positionner, en calculant la *proportion de communes* qui ont moins de N habitants : c'est la notion de fréquence (cumulée) qui apparaît ici pertinente. Il y a 35868 communes, et donc $3391/35868 \approx 9,5$ % des communes qui ont moins de 100 habitants.

On peut aussi regarder quelle *proportion de la population totale* comptent les communes de moins de N habitants, toujours en prenant quelques valeurs de N . Si, par

⁴ Dans un premier temps, on leur avait demandé de « préciser comment il faudrait faire pour savoir si une ville de 7000 habitants, c'est une ville qui a beaucoup d'habitants ou une ville qui n'a pas beaucoup d'habitants. »

⁵ On notera que la ressource thématique susmentionnée indique : « Les notions de fréquences et de fréquences cumulées, de moyenne et de médiane sont introduites dès la classe de 5e, celle d'étendue en classe de 4e, dans des situations qui leur donnent du sens. » (MEN, 2016, p. 2)

exemple, on considère les communes de moins de 1000 habitants, on constate qu'elles rassemblent 14 % de la population française tout en représentant $26042/35868 \approx 73$ % des communes françaises. Il faudra considérer les communes de moins de 400 habitants pour constater qu'elles rassemblent 5 % de la population française tout en représentant $16015/35868 \approx 47$ % des communes françaises : on peut convenir que ce sont des petites communes.

Nous reproduisons ci-dessous les propositions de deux équipes, D et E. L'équipe D (voir figure 3) propose une réponse construite, mais qui s'appuie sur deux notions, *moyenne* et *médiane*, permettant uniquement de démarrer l'enquête, avant de mentionner que la détermination des quartiles (et des déciles) permettrait d'« affiner cette réponse ». Le calcul des fréquences cumulées, qui relève du programme du cycle 4 (voir note 3) et s'avère ici très pertinent, n'est pas même mentionné. La proposition de l'équipe E (voir figure 4) est caricaturale : le tableau fournit la liste des indicateurs (de position et de dispersion) que semblent connaître les membres de l'équipe. Notons tout de même que la dernière ligne donne la fréquence des villes de plus de 7 000 habitants (mais rien n'est mentionné sur l'obtention de ce résultat).

Figure 3

Proposition de l'équipe D

Qu'est-ce qu'une commune française qui n'a pas beaucoup d'habitants ?
Pour déterminer d'abord de façon large à quelle moitié (supérieure ou inférieure) appartient une commune (en termes de nombre d'habitants), il convient de commencer par calculer la moyenne et déterminer la médiane de la série statistique. La moyenne est de 1798,4 (à 0,1 près) et la médiane est 437. On peut donc dire qu'en dessous de 437 habitants, la commune fait partie des 50 % de communes ayant le moins d'habitants.
Pour affiner cette réponse, nous pouvons calculer les quartiles. On s'aperçoit au vu des tableaux que le premier quartile vaut : 196. (hors programme explicite du collège mais peut-être introduit à cette occasion, de même que les déciles).
On peut donc dire qu'une ville avec moins de 196 habitants est une petite ville.

Figure 4
Proposition de l'équipe E

Moyenne (population)	1789
Mediane	437
QUARTILES	
Q1	196
Q2	437
Q3	1083
DECILES	
D1	102
D2	108
D3	169
D4	239
D5	327
D6	451
D7	628
D8	909
D9	1467
D10	3195
Ecart-type	
	2205
Fréquence villes +7000hab	
	0,03943

Pour compléter notre propos, considérons maintenant quelques-unes des remarques faites par l'ensemble des équipes sur la consigne proposée :

- On peut proposer un tel sujet à travailler en classe à des élèves de collège, d'abord en petit groupe puis à débattre en classe entière, à condition d'avoir préparé en amont les questions de relance. Maintenant, avec le programme de collège où il n'y a plus les quartiles, le sujet nous semble loin des attentes actuelles.
- Les élèves ne vont-ils pas perdre du temps à regarder le nombre d'habitants dans certaines villes ?
- Le fichier n'est-il pas trop imposant au niveau des données pour un élève ?

Ces deux propositions et ces remarques permettent de dégager un certain nombre d'obstacles rencontrés par les élèves professeurs. Dans leur ensemble, les équipes se focalisent d'emblée sur les éléments technologiques les plus « savants » mathématiquement parlant (moyenne, médiane, quartiles, déciles, etc.), qui constituent autant d'éléments attracteurs. Aucune équipe n'a évoqué le traitement des données ni le calcul des fréquences et des fréquences cumulées. On notera à ce sujet la difficulté qu'ont eue les élèves professeurs à traiter les données⁶ : pour un certain nombre d'entre eux, cela a été l'occasion de rencontrer la fonction « Trier » du tableur.

⁶ La taille des données (communes de France) est adaptée à des élèves professeurs ; cette enquête a déjà été proposée dans des classes de collège, mais à l'échelle d'un département.

Proposer des situations riches comme celle des « petites communes » peut conduire la classe à prendre des chemins variés, et le rôle de directeur d'étude est alors crucial pour faire avancer l'étude. Or la *gestion de l'étude*, dont la maîtrise constitue un objectif essentiel dans leur formation, reste d'une grande difficulté pour les professeurs stagiaires⁷ et ils ont l'impression qu'en proposant des activités très fermées et guidées, ils parviennent à avoir une plus grande maîtrise du temps et de l'avancée de l'étude, et par ce biais une meilleure *gestion de classe*. Cette dernière est bien souvent la principale préoccupation des professeurs stagiaires et il est très difficile pour eux d'entendre que cette gestion est très étroitement liée à la robustesse des activités qu'ils proposent et à leur gestion de l'étude, avec la relance par le biais des questions cruciales.

La remarque « Les élèves ne vont-ils pas perdre du temps à regarder le nombre d'habitants dans certaines villes ? » est révélatrice de certaines contraintes qui pèsent sur la formation : chaque activité doit conduire à un seul objectif et il n'y a aucune place pour l'émergence de plusieurs organisations mathématiques ponctuelles et leur amalgamation. Il est vrai que de ce point de vue, si chaque organisation mathématique ponctuelle doit donner lieu à une activité, le temps de l'horloge va défiler très vite ! Or, dans le cas que nous venons de présenter, quand on détermine, par exemple, le nombre d'habitants de la ville où on habite, on travaille sur le type de tâches « Lire des données sous forme de données brutes » et on consolide ses « capacités à prélever de l'information sur divers supports » (MEN, 2016, p. 2).

Cette étude statistique a aussi été l'occasion de mettre en avant un certain nombre de manques chez les élèves professeurs. On constate par exemple qu'ils ont beaucoup de difficultés à travailler avec un grand nombre de données. Derrière des remarques du style :

⁷ Très majoritairement, les élèves professeurs qui sont en deuxième année de formation à l'ESPE sont *professeurs stagiaires* : il s'agit pour eux d'une formation en alternance, la moitié de leur temps de travail étant consacré à un *stage* dans un établissement scolaire (collège ou lycée), l'autre moitié à la formation à l'ESPE.

« C'est impossible de traiter autant de données ! », « Les élèves n'y arriveront pas », c'était leur propre difficulté avec l'outil tableur qu'ils dissimulaient derrière les compétences de leurs élèves ; on retrouve ici ce que Gisèle Cirade (2006, p. 188) appelle une « difficulté-écran ». Par ailleurs, les gestes propres à l'enquête (Chevallard & Ladage, 2011) sont absents de l'équipement praxéologique des élèves professeurs : dans le cas de l'enquête sur les petites communes, *a priori* aucun groupe n'a consulté la notice accompagnant le fichier sur le site de l'INSEE ou cherché un article à propos de la définition d'une « petite commune », d'une « grande commune ».

Contraintes liées aux mathématiques

Revenons au secteur « Organisation et gestion de données ». Les élèves professeurs ont eu beaucoup mal à développer l'analyse de l'organisation mathématique et il faut dire que, même si ce secteur occupe une place non négligeable dans les programmes du secondaire, il est considéré comme non prioritaire : dans les progressions, ce qui relève de la statistique descriptive est bien souvent traité en fin d'année, « si on a le temps »⁸. Il est vrai qu'un professeur est amené à traiter des données tout au long de l'année, mais l'organisation mathématique propre à l'étude d'une série statistique est rarement enjeu de l'étude. Pour la plupart, tout au long de leur cursus universitaire, les élèves professeurs actuellement en formation n'ont jamais eu de cours de statistique, ou seulement par le biais d'une option pour appréhender des concepts plus théoriques tels que ceux que l'on rencontre en statistique inférentielle pour la validation de modèles. On retrouve là un phénomène analogue à celui qui a été identifié par Yves Chevallard et Floriane Wozniak (2007), qui disent que « si la réduction arithmétique de la statistique est écologiquement possible, elle n'est guère durable, en ce sens que, se situant à un bas

⁸ Sur un échantillon de 16 progressions communes dans des classes de cycle 4 de l'académie de Toulouse, la séquence « Statistiques » (c'est ainsi qu'elle est nommée) est abordée pour 70 % d'entre elles au troisième trimestre.

niveau mathématique, elle est peu significative d'une discipline spécifique supposée, 'la statistique' » (p. 10). Ils poursuivent en donnant un exemple simple, mais éclairant :

Si l'élève doit trouver la moyenne annuelle de Fanny, qui a eu 6,9 de moyenne au premier trimestre, 9,8 de moyenne au deuxième trimestre et 9,3 de moyenne au troisième trimestre, il peut effectuer mentalement le calcul suivant : $6,9 + 9,8 + 9,3 = (6 + 0,9 + 0,8 + 0,3) + 9 + 9 = (6 + 1,7 + 0,3) + 9 + 9 = 8 + 9 + 9$. Or ce calcul, loin d'être emblématique d'une science « à forte personnalité », relève en vérité de l'arithmétique scolaire la plus banale qui soit ! N'ayant pas de spécificité, n'étant pas significatif de quelque chose de spécifique, le « travail mathématique » correspondant tendra à se réduire à la portion congrue, voire à être escamoté : le fait de donner ce travail « en DM » ou de renvoyer les élèves au manuel pour qu'ils s'aident eux-mêmes à ce propos sont les effets les plus communs de cette dépréciation scolaire. (Ibid., p. 10)

La statistique descriptive, qui n'est pas enseignée dans les cursus universitaires tout en étant essentielle pour les mathématiques enseignées au secondaire, pourrait être qualifiée de savoir oublié. Dans le cas que nous considérons, on peut ajouter que certains éléments technologiques sont « non stabilisés » : une étude des manuels du cycle 4 met en évidence un flou persistant concernant certaines notions (données, valeur, caractère, série), comme nous allons le voir.

À propos de la notion d'effectif

Considérons tout d'abord la notion d'effectif. On reproduit ci-dessous quatre définitions proposées dans des manuels⁹ de cycle 4 :

[C4, Delta]. Dans une série de données, l'**effectif d'une** donnée est le nombre de fois où elle apparaît. L'**effectif total** est le nombre de valeurs de la série statistique. (Lambotte, 2016, p. 208)

[5e, Indigo]. Dans une série de données : L'effectif d'une donnée est le nombre de fois où cette donnée apparaît. L'effectif total est la somme de tous les effectifs. (Barnet, 2016, p. 152)

[C4, Transmath]. Lors d'une enquête, une liste de données a été relevée. L'**effectif** d'une donnée est le nombre de fois où cette donnée apparaît dans la liste. L'**effectif total** est le nombre total de données dans la liste. (Carlod, 2016, p. 194)

⁹ Les manuels sont repérés entre crochets par le *niveau* (C4 pour cycle 4 – élèves de 12-15 ans – ou l'un des trois niveaux de classe : 5e, 4e, 3e) et la *collection*.

[C4, Kiwi]. Dans une série statistique, l'**effectif** d'une valeur est le nombre de fois où cette valeur apparaît dans la série. L'**effectif total** est la somme de tous les effectifs. (Beltramone, 2016, p. 68)

Sur ces quatre manuels, trois définissent l'effectif d'une *donnée* et un l'effectif d'une *valeur*. Mais qu'est-ce qu'une donnée ? Qu'est-ce qu'une valeur ? Aucun manuel ne propose de poser, même de façon informelle, les premiers éléments du modèle classiquement utilisé en statistique (population, individu, caractère, valeur) avant d'énoncer les notions enjeu de l'étude (effectif, fréquence, etc.). Revenons à la définition d'effectif, en nous appuyant sur un ouvrage rédigé par Gérard Chauvat et Jean-Philippe Réau (1995). On y lit notamment que « l'**effectif total** de la population est le nombre N d'individus qui la constitue » (*ibid*, p. 10), que, « dans la pratique, la contingence est constituée des observations, c'est-à-dire des valeurs observées du (ou des) caractère(s) de la population au(x)quel(s) on s'intéresse » (*ibid*, p. 11) et qu'« on appelle **effectif** de la valeur v_i le nombre de fois où elle a été observée » (*ibid*, p. 13).

Dans le manuel [C4, Delta], une *donnée* correspond visiblement à ce qu'on appelle classiquement une *valeur* et l'effectif total est défini comme étant le nombre de *valeurs*, alors qu'il s'agit ici des *valeurs observées*. Le manuel [5e, Indigo] donne une définition de l'*effectif total* qui le rend tributaire de celle des effectifs des valeurs et qui, en fait, renvoie plutôt à une propriété. Le manuel [C4, Transmath] impose que les données soient fournies sous forme d'une liste : que faire dans le cas où l'on connaît la distribution des observations ? Ces différents flous dans les définitions des manuels confirment une certaine *péjoration* du savoir statistique.

À propos de la notion de médiane

Poursuivons notre enquête en nous intéressant à la médiane. Nous reproduisons ci-dessous deux définitions proposées dans des manuels de cycle 4 :

[C4, Maths Monde]. La médiane d'une série statistique est un nombre tel qu'il y ait autant de valeurs inférieures ou égales à ce nombre que de valeurs supérieures ou égales à ce nombre. (Lanata, 2016, p. 71)

[C4, Delta]. La médiane d'une série de valeurs est une valeur telle que : au moins 50 % des valeurs lui soient inférieures ou égales ; au moins 50 % des valeurs lui soient supérieures ou égales. (Lambotte, 2016, p. 209)

Essayons alors de déterminer la médiane de la série suivante¹⁰ : (1 ; 2 ; 2 ; 2 ; 2).

Dans le cas de la première définition, on peut montrer qu'il n'existe pas de nombre m tel qu'il existe *autant* de valeurs observées $\leq m$ que de valeurs observées $\geq m$. Dans le cas de la seconde définition, on peut montrer qu'il existe bien une valeur m qui convient : $m = 2$. Mais si l'on utilise cette définition dans le cas de la liste (1 ; 2 ; 3 ; 4 ; 5 ; 6), on peut montrer que les valeurs $m = 3$ et $m = 4$ conviennent, alors que le manuel propose de définir la médiane d'une série de valeurs. La situation n'est pas simple, d'autant que les définitions qui sont proposées ne sont pas des plus opérationnelles.

Les documents foisonnent sur la médiane. On peut par exemple citer un texte proposé par l'inspection pédagogique régionale de mathématiques de l'académie de Strasbourg, qui fait le point sur « les difficultés soulevées par la définition de la médiane et des quantiles », en « classes de premières L, ES, S » (Meyer, s.d.). On peut aussi noter qu'une ressource produite dans le cadre de l'accompagnement des programmes de mathématiques publiés en 2008 propose une définition exploitable (MEN, 2007, p. 11) :

La définition qui est retenue en collège pour la médiane d'une série est celle qui est adoptée dans le programme de seconde. Elle s'appuie sur la pratique :

Médiane (empirique) : *La série des données est ordonnée par ordre croissant. Si la série est de taille impaire ($2n+1$), la médiane est la valeur du terme de rang $n+1$. Si la série est de taille paire ($2n$), la médiane est la demi-somme des valeurs des termes de rang n et $n+1$.*

D'autres définitions sont parfois utilisées ; par exemple, la médiane est le deuxième quartile.

L'enseignement des quartiles relevant depuis 2016 du programme de la classe de seconde, cette autre définition (la médiane est le deuxième quartile) n'entre pas en

¹⁰ Pour simplifier, on ne considère qu'un petit nombre d'observations, avec les valeurs observées déjà rangées dans l'ordre croissant

concurrence avec la définition empirique énoncée plus haut au cycle 4, sauf peut-être par le biais des calculatrices et des tableurs, ainsi qu'en témoigne un article posté par Yohan Cecere (2017) sur le blog Kwyk Actu : dans un environnement numérique anglo-saxon, la médiane est définie comme étant le deuxième quartile d'une série. Des élèves professeurs ont été déstabilisés lorsque des calculatrices dans la classe ont donné des réponses différentes pour la médiane. On retrouve un phénomène déjà repéré : les élèves professeurs ont des difficultés pour mener une enquête suffisamment approfondie autour de l'organisation mathématique.

Conclusion

Le travail sur l'enseignement du secteur « Organisation et gestion de données » au cycle 4 que nous venons de présenter permet de révéler des *difficultés récurrentes* rencontrées par les professeurs dans l'exercice de leur métier, même si l'enquête que nous avons menée est partie de celles qui étaient rencontrées par les élèves professeurs, avec notamment la mise en évidence de difficultés « écran » chez nombre d'entre eux. Nous avons mis en évidence une *péjoration* du savoir lié à ce secteur : on a là un exemple de la difficulté d'existence d'un rapport institutionnel aux objets de la statistique pour la position de professeur prenant en charge la variabilité, ce qui conduit à une instabilité de certaines notions et à de nombreuses imprécisions ou inadéquations, les définitions proposées dans les manuels relevant parfois de la technique ou n'étant pas du tout opérationnelles. Ce savoir, pourtant stabilisé depuis longtemps, demanderait à être retravaillé en vue de son enseignement dans le secondaire. Comme nous le disions plus haut, on est face à un savoir qui nous semble *oublié*.

Références

- Barnet, C. (Éd.) *Mathématiques, cycle 4, 5^e*. Collection Mission Indigo, Paris : Hachette, 2016.
- Beltramone, J.-P. (Éd.) *Mathématiques, cycle 4*. Collection Kiwi, Paris : Hachette, 2016.
- Carlod, V. (Éd.) *Mathématiques, cycle 4*. Collection Transmath, Paris : Nathan, 2016.
- Cecere, Y. *Quartiles et médiane : pourquoi ma calculatrice se trompe ?* 27 juin 2017
<http://blog.kwyk.fr/2017/06/quartiles-et-mediane-pourquoi-ma.html>
- Chauvat, G. & Reau, J.-P. *Statistique descriptive*. Paris : Hachette, 1995.
- Chevallard, Y. Le passage de l'arithmétique à l'algèbre dans l'enseignement des mathématiques au collège. Perspectives curriculaires : la notion de modélisation. *Petit x*, 19, p. 43-72, 1989.
- Chevallard, Y. Passé et présent de la théorie anthropologique du didactique. In : *Sociedad, escuela y matemáticas. Aportaciones de la teoría antropológica de lo didáctico* (TAD), Jaén, Espagne : Publicaciones de la Universidad de Jaén, p. 705-746, 2007.
- Chevallard, Y. L'évolution du paradigme scolaire et le devenir des mathématiques. Questions vives et problèmes cruciaux. In : *Questions vives en didactique des mathématiques : problèmes de la profession d'enseignant, rôle du langage*, Grenoble : La pensée Sauvage, p. 85-120, 2013.
- Chevallard, Y. & Ladage, C. Enquêter avec l'Internet. Études pour une didactique de l'enquête. *Éducation & didactique*, 5(2), p. 85-116., 2011
- Chevallard, Y. & Wozniak, F. Enseigner la statistique : un problème de la profession. In : *La statistique dans l'enseignement et la formation PLC2 & Les résistances et les changements dans les pratiques d'enseignement en formation initiale*, p. 8-25, 2007. <http://www.univ-irem.fr/spip.php?rubrique442>
- Cirade, G. Devenir professeur de mathématiques : entre problèmes de la profession et formation en IUFM. Les mathématiques comme problème professionnel, Thèse de doctorat), 2006. <https://tel.archives-ouvertes.fr/tel-00120709>
- LAMBOTTE, L. (Éd.). *Mathématiques, cycle 4*. Collection Delta, Paris : Belin, 2016.
- Lanata, F. (Éd.) *Mathématiques, cycle 4*. Collection Maths Monde). Paris : Didier, 2016.
- Meyer, E. *Les difficultés soulevées par la définition de la médiane et des quantiles. Classes de premières L, ES, S*, sans date. https://www.ac-strasbourg.fr/fileadmin/pedagogie/mathematiques/Inspection/Documentations/m-edquan_E_Meyer.doc
- Ministère De L'éducation Nationale. *Enseigner au collège. Mathématiques. Programmes et accompagnement*. Paris : CNDP, 2004.
- Ministère De L'éducation Nationale. Organisation et gestion de données au collège. In : *Ressources pour les classes de 6e, 5e, 4e, et 3e du collège*, 2007. http://cache.media.eduscol.education.fr/file/Programmes/17/5/doc_acc_clg_organisation_donnees_109175.pdf
- Ministère de L'éducation Nationale. Programmes pour les cycles 2, 3, 4. *Bulletin officiel spécial n°11 du 26 novembre 2015*, 2015

Ministère de L'éducation Nationale. Interpréter, représenter et traiter des données. In : *Ressources d'accompagnement du programme de mathématiques (cycle 4)*, 2016. http://cache.media.eduscol.education.fr/file/Traitement_des_donnees/03/6/RA16_C4_MATH_doc_maitre_564036.pdf

Annexe 1. Production de l'équipe A

Activité 1 : (Créé de toutes pièces)

Nous nous intéresserons à l'étude des pointures de chaussures au sein de la classe.

Etape 1 : Recensons les pointures existantes au sein de la classe dans le tableur suivant.

NB : si classe d'effectif pair : au départ, le professeur ajoute sa pointure.

NOM	Elève 1	Elève 2	Elève 3
Pointure	37	35	39

Etape 2 : Rangeons ces pointures par ordre croissant.

NOM	Elève 2	Elève i	Elève j
Pointure	35	35	36

Activité 1 : (Créé de toutes pièces)

Nous nous intéresserons à l'étude des pointures de chaussures au sein de la classe.

Etape 1 : Recensons les pointures existantes au sein de la classe dans le tableur suivant.

NB : si classe d'effectif pair : au départ, le professeur ajoute sa pointure.

NOM	Elève 1	Elève 2	Elève 3
Pointure	37	35	39

Etape 2 : Rangeons ces pointures par ordre croissant.

NOM	Elève 2	Elève i	Elève j
Pointure	35	35	36

Annexe 2. Production de l'équipe B

Activité :

L'organisateur d'une compétition amicale de judo souhaite répartir les combattants en deux poules. On souhaite que cette répartition soit la plus équitable possible. La première poule est dite des « légers » et la seconde est dite des « lourds ».

On a donc relevé le poids en kg de chacun des judokas :

$$62 - 98 - 78 - 95 - 68 - 59 - 74 - 81 - 102 - 71 - 80 - 61 - 65 - 72 - 65$$

- 1) Peut-il y avoir autant de judokas dans chacune des deux poules ?
- 2) a) Calculer le poids moyen des judokas.
Peut-on utiliser ce poids pour répartir les judokas dans les deux poules ?
b) Quel est le poids du judoka que l'on peut placer indifféremment chez les « lourds » ou chez les « légers » ?
- 3) Le compétiteur le plus lourd (102 kg) décide de se retirer de la compétition.
a) Combien y aura-t-il de judokas dans chacune des deux poules ?
b) Quels poids limites peut-on donner pour définir la catégorie des lourds et celle des légers ?

Annexe 3. Production de l'équipe C

Modification de l'énoncé

- Voici la liste des tailles, en Mo, des vidéos que Claire a prises avec son téléphone portable. Déterminer la valeur de la série telle qu'au moins 50% des valeurs de la série lui soient inférieur.

(Question d'aide : ranger par ordre croissant la liste)

485 Mo	480 Mo	520 Mo	430 Mo
270 Mo	380 Mo	295 Mo	385 Mo
415 Mo	390 Mo	280 Mo	

- Le tableau ci-dessous présente la répartition des tailles des vidéos dans le téléphone de Karl.
 - Calculer l'effectif cumulé croissant ?
 - Déterminer la valeur de la série telle qu'au moins 50% des valeurs de la série lui soient inférieur.

Taille des vidéos (en Mo)	150	160	180	210	230	240	280
Effectif	2	5	7	4	3	2	4

Modification de l'énoncé

- Voici la liste des tailles, en Mo, des vidéos que Claire a prises avec son téléphone portable. Déterminer la valeur de la série telle qu'au moins 50% des valeurs de la série lui soient inférieur.

(Question d'aide : ranger par ordre croissant la liste)

485 Mo	480 Mo	520 Mo	430 Mo
270 Mo	380 Mo	295 Mo	385 Mo
415 Mo	390 Mo	280 Mo	

- Le tableau ci-dessous présente la répartition des tailles des vidéos dans le téléphone de Karl.
 - Calculer l'effectif cumulé croissant ?
 - Déterminer la valeur de la série telle qu'au moins 50% des valeurs de la série lui soient inférieur.

Taille des vidéos (en Mo)	150	160	180	210	230	240	280
Effectif	2	5	7	4	3	2	4