L'ENSEIGNEMENT DE LA STATISTIQUE : OÙ, QUAND, COMMENT, POURQUOI PAS ?

Linda GATTUSO[1]

TITLE

The teaching of statistics: where, when, how and why not?

RÉSUMÉ

L'importance de la statistique est largement admise et son utilisation est répandue. Pour éviter son emploi inadéquat sinon erroné, il serait nécessaire d'inclure son enseignement dans la formation, mais où, quand et comment ? Toutes ces questions sont interreliées. Où : dans un contexte de transfert des compétences, il n'est pas exclu que la statistique soit présente en géographie, en économie ou encore en physique par exemple, là où l'on trouve des données. Toutefois, le plus souvent, s'il se trouve à l'école, l'enseignement de la statistique s'insère dans le cours de mathématiques. Quels sont les avantages et inconvénients de cette situation ? Quand : à quel moment peut-on le faire, à l'école ou vaudrait-il mieux attendre l'université ? Quand les élèves sont-ils prêts à comprendre les concepts de la statistique ? Quel que soit l'endroit ou le moment, il reste qu'il faut se demander comment cet enseignement doit être fait. Un enseignement de la statistique favorisant le développement de la pensée statistique et la construction de concepts de base comme ceux de variabilité, de distribution, peut commencer sans l'utilisation *a priori* de développements mathématiques formels. Le développement de la pensée et du raisonnement statistique passe par l'expérimentation du parcours du statisticien, en bref, questionnement, collecte de données, analyse et conclusion. De plus, l'apprentissage doit se faire en tenant compte des acquis (plus ou moins conscients) de l'apprenant. Reste la question du « Pourquoi pas ? ». Quels sont aujourd'hui les obstacles freinant l'essor de l'enseignement de la statistique ? Nous tenterons de répondre à toutes ces questions tout en apportant quelques suggestions de solution.

Mots-clés : *enseignement de la statistique, enseignants, formation.*

ABSTRACT

The importance of statistics is well known and its use is widely spread. To avoid its inadequate if not erroneous use, it would be necessary to include the teaching of statistics in the school curriculum, but where, when and how? All these questions are related. Where? In a context of transfer of competencies, it is not excluded that statistics should be present in geography, in economics or in physic for example, where the data are. However, more often, if in school, the teaching of statistics is done in the mathematic class. What are the advantages and inconveniences of this situation? When? Can it be done in school or would it be better to wait until university? When are pupils ready to understand statistical concepts? Whatever the place or the time, the question remains how it should be done. A teaching of statistics favouring the development of statistical thinking and the construction of basic concepts such as variability and distribution can begin without the use *a priori* of formal mathematical developments. The development of statistical thinking and reasoning goes through the experience of the practice of the statisticians, in brief, questioning, data collection, analysis and conclusion. Furthermore, the learning must take into account prior experience (more or less conscious) of the learner. But the question remains: « why not? ». What are the obstacles that slow down the progress of the teaching of statistics? We will try to answer all these questions while suggesting some solutions.

Keywords: *teaching of statistics, teachers, training.*

[1] Université du Québec à Montréal, gattuso.linda@uqam.ca

Statistique et Enseignement, **2**(1), 5-30, http://www.statistique-et-enseignement.fr/ojs/

1 Introduction

Pourquoi l'enseignement de la statistique ? La nécessité pour un citoyen d'aujourd'hui de comprendre et d'interpréter la statistique qui est présente de plus en plus dans les médias n'est plus à démontrer. Aujourd'hui, il est difficile de participer aux différents débats concernant l'éducation, la santé, l'environnement, etc., sans comprendre et interpréter des données et des graphiques statistiques, et sans juger de la validité des résultats (Konold & Higgins, 2003 ; Cobb, 1999). Ce fait étant dorénavant largement accepté, plusieurs pays ont inclus la statistique dans le curriculum scolaire. Cependant, des questions demeurent : où, quand, comment et pourquoi pas ?

Les réponses proposées à ces questions découlent à la fois de mon expérience professionnelle de plus de 20 ans d'enseignement de la statistique de base, complétée par une carrière de didacticienne concentrée sur la problématique de l'enseignement de la statistique.

C'est dans un cégep québécois (17-18 ans) que j'ai enseigné la statistique. À l'époque, il y avait dans les divers programmes au moins 10 cours différents de statistique qui étaient soit généraux ou plus spécialisés comme, par exemple, le « Contrôle de qualité » pour des étudiants de technique de génie civil. Le niveau des divers cours différait aussi, certains exigeaient des préalables de calcul différentiel[2], d'autres pas.

À l'origine (1969), le cours de statistique était très théorique, avec de l'analyse combinatoire à profusion (photos claires ou pas, équipe de volley ou de hockey, etc.), des probabilités traditionnelles, des statistiques descriptives et des lois de probabilités de base : binomiale, poisson et normale sans vraiment d'exemples réels. Que pouvaient faire les étudiants avec ça ?

J'ai vite compris qu'il fallait « concrétiser » ces cours et amener les étudiants à une connaissance plus applicable de la statistique. Les calculatrices programmables sont arrivées vers la mi-80 et plus tard, les ordinateurs (au début très limités), outils technologiques qui ont multiplié les possibilités d'activités, d'expérimentations, etc. Il y a donc eu plusieurs années d'expérimentations qui ont été très instructives et qui pour certaines ont porté fruit.

Il va sans dire que les idées exprimées ici sont largement partagées dans le milieu scientifique de l'enseignement de la statistique. C'est pourquoi j'appuierai la plupart de mes propos sur des auteurs que je citerai quelquefois au passage, sinon en référence, à la fin de l'article.

Une dernière précision : il sera ici question de l'enseignement de la statistique à l'école et non dans un cadre universitaire, bien que certaines propositions pourraient être facilement utilisées dans les cours de base universitaires offerts aux non spécialistes.

Toutes les questions présentées sont interreliées mais, puisqu'il faut établir un ordre d'entrée, commençons par le « où ? ».

2 L'enseignement de la statistique : où ?

Il y a lieu de se demander si la statistique doit être enseignée dans le cadre du cours de mathématiques ou intégrée dans diverses disciplines qui en font l'utilisation. À l'heure de la

[2] Entre autres pour définir les fonctions de densité d'une variable continue.

transversalité, on pourrait insérer l'enseignement de la statistique dans plusieurs matières scolaires. En géographie, on peut comparer la longueur des fleuves, étudier les distributions des populations selon l'âge ; en sciences, on peut examiner plus systématiquement certaines expérimentations, regarder les diverses composantes de plusieurs types de céréales de déjeuner : protéines, glucides, etc., et bien plus encore. Et cela devrait être fait ! Ce serait l'occasion de souligner l'utilisation et les applications de la statistique dans divers domaines scientifiques et dans la vie quotidienne. Il va de soi qu'il est nécessaire de s'assurer que l'introduction d'éléments de la statistique dans les cours de diverses disciplines soit présentée correctement par des enseignants ayant la formation nécessaire.

Il reste qu'il y a là un danger que la statistique soit avalée par la discipline hôte (biométrie, économie, informatique, etc.) (Moore & Cobb, 2000 ; Bessant & MacPherson, 2002) et perde ainsi son identité propre. Pour cette raison notamment, je crois, comme d'autres (Schield, 2001 ; Scheaffer, 2002), que dans les conditions actuelles, la statistique devrait être enseignée dans les cours de mathématiques. La statistique enseignée dans le cours de mathématiques peut ainsi « rester un tout » et garder une entité propre (un peu comme la géométrie) au lieu d'être parsemée ou même diluée. On peut le voir comme un avantage. De plus, il faut être réaliste, pour le moment, il n'y a pas assez de statisticiens ou d'enseignants de statistique pour prendre en charge son enseignement. Il faudra sûrement y penser et prévoir une formation adéquate pour les enseignants.

Actuellement, là où la statistique apparaît dans les programmes scolaires, elle est presque toujours insérée dans le programme de mathématiques. Et quand on arrive au niveau collégial (collège américain ou cégep québécois), c'est dans les départements de mathématiques que l'on retrouve le plus grand nombre d'étudiants qui suivent des cours de statistique. Cependant, il faut être conscient que trop souvent ces cours sont donnés par des professeurs dont le principal intérêt est ailleurs. Leur intérêt et leur formation sont presque toujours plutôt du côté de la mathématique.

Pour le moment, c'est un fait avec lequel il faut composer. Il est donc très important de mettre en évidence la différence entre la statistique et la mathématique, car différence il y a et les enseignants doivent la connaître.

Shaughnessy dit même que « …la statistique n'est pas la mathématique, même pas une branche de la mathématique… » (2006), et Moore et Cobb rajoutent : « …bien que la statistique est une science mathématique, ce n'est pas un sous-domaine de la mathématique… » (2000).

3 Différences, similitudes et tensions

Bien qu'il y ait entre l'enseignement de la mathématique et de la statistique plusieurs similitudes, les différences sont importantes et parfois cause de tensions.

3.1 Le contexte

En analyse de données[3], le contexte donne un sens. On ne peut parler de nombres sans les relier au contexte duquel ils sont tirés, et on ne connaît pas le sens ou l'intérêt des nombres

[3] Dans cet article, « Analyse de données » doit être compris dans le sens utilisé dans le monde anglophone. Cobb et Moore (2000) disent que l'analyse des données est la forme contemporaine de la « statistique

tant que l'on ne connaît pas le contexte. Les nombres seuls ne peuvent donner lieu à une analyse de données, soulignent Cobb et Moore (*id.*). La statistique demande une autre sorte de pensée parce que les nombres ne sont pas seulement des nombres mais des nombres avec un contexte.

Même si les mathématiciens font souvent appel au contexte pour motiver ou comme source de problèmes de recherche, la raison première de la pensée mathématique est le modèle abstrait. À l'école, on étudie les nombres et leurs opérations sans qu'il soit absolument nécessaire de les mettre en contexte. De plus, Meletiou (2003) affirme que trop souvent les enseignants considèrent que la « bonne » mathématique est pure et « défaite » du contexte.

3.2 La modélisation et les données

Dans l'enseignement des deux disciplines, mathématique et statistique, il est question de modélisation et cela pourrait servir de pont entre les deux. Cependant, selon Biehler (2008), dans la classe de mathématique, la modélisation ne porte pas attention à la collecte et à l'analyse des données qui sont nécessaires si on se place dans une perspective statistique. En mathématique, on recherchera les modèles pour abstraire, généraliser et générer des lois et des algorithmes (Shaughnessy, 2006).

Traditionnellement, dans la classe de mathématique, on présente une structure terminée et les exemples viennent ensuite. Les élèves reçoivent un produit fini et ne participent pas à la procédure de construction. On néglige la variabilité qui est omniprésente dans le réel et qui est un concept clé de la statistique. Les données présentées ne sont pas réelles, ni même réalistes, et s'appliquent « trop » au modèle. Les points sont à peu près tous sur la droite. On néglige même de dire que le modèle obtenu est un *modèle* de la réalité. On utilise des exemples pseudo-réels, comme la formule « vitesse = distance/temps » et on ne tient pas compte de la variabilité qu'auraient les données si elles étaient recueillies expérimentalement (Meletiou, 2003 ; Moore & Cobb, 2000 ; Bieheler, 2008).

La résolution de problèmes réels où la source des données est explicitée (à défaut d'en faire la cueillette), où les variations ne sont pas estompées et où l'importance du contexte n'est pas négligée, serait avantageuse à la fois pour la statistique et la mathématique.

3.3 Incertitude

En statistique, la solution est pour ainsi dire indéterminée. Elle dépendra sûrement du contexte, mais aussi de toute la procédure : les hypothèses de départ, les choix dans la planification et l'analyse, etc.

Selon Garfield (1999), les élèves voient ce qu'ils apprennent à l'école comme prédéterminé et réglé car les problèmes de mathématiques ont tendance à avoir une seule

descriptive » renforcée par des outils plus nombreux et plus élaborés et, plus particulièrement, par une philosophie due en grande partie à Tukey. Les termes « analyse exploratoire des données » (EDA, exploratory data analysis) reflètent cette philosophie. Cette analyse des données est un prérequis à l'inférence formelle. Dans le wikipedia de langue anglaise : "Analysis of data is a process of inspecting, cleaning, transforming, and modeling data with the goal of highlighting useful information, suggesting conclusions, and supporting decision making". Traduction de l'auteur : l'analyse des données est un processus d'inspection, de nettoyage, de transformation et de modélisation des données ayant comme but de soulever des informations utiles, de suggérer des conclusions et de soutenir la prise de décision.

bonne solution. En étudiant l'analyse des données et la statistique, ils peuvent constater que les solutions à certains problèmes dépendent des hypothèses et sont incertaines. Un problème d'analyse de données n'a pas dans la plupart des cas de solution prédéterminée, mais plutôt plusieurs raisonnables.

Par ailleurs, dans la classe de mathématique, on habitue les élèves à rechercher des modèles, par exemple, des suites de nombres. Alors, ils ont tendance à faire cela aussi lorsqu'on leur demande, par exemple, des résultats possibles pour des lancers de 10 pièces de monnaies: « A peu près autant de piles que de faces, il faut à peu près 5P et 5F ». Les élèves prédisent l'incertain selon un modèle prédéterminé (Rubel, 2006 ; Meletiou, 2003).

En mathématique, les théorèmes, une fois prouvés, sont vrais ; une preuve bien faite est « éternelle ». Or, il n'y a rien comme ça en statistique ; les méthodes sont parfois efficaces si elles sont utilisées avec habileté et justesse, mais il n'y a pas de certitude. Les résultats dépendent de l'exactitude du travail (Moore & Cobb, 2000). C'est pourquoi il faut garder une attitude critique et se questionner sur tout, la question, les données et leur cueillette, la méthode d'analyse, etc. Il faudra, dans la classe de statistique, habituer les élèves à questionner toutes les étapes du travail : une pratique qui peut également être profitable en mathématique.

3.4 Mesure

Mentionnons rapidement un autre élément qui présente des distinctions selon qu'on se place du point de vue de la mathématique ou de la statistique : la mesure. C'est une notion fondamentale de la mathématique à l'école : longueurs, aires et volumes en sont les exemples types. Dans ce cas, bien que l'on apprenne aussi à manipuler des outils de mesurage comme la règle par exemple, l'importance de l'exactitude de la mesure est relative. En effet, la précision de la figure, en géométrie par exemple, est secondaire à l'adoption d'hypothèses justes : « Soit la mesure de $\overline{AB}$ égale à la mesure de $\overline{BC}$… ». Or, en statistique, il faut concevoir de nouvelles procédures de mesure. Comment mesurer qui n'a pas d'emploi, la taille de la famille, le nombre de maisons sur la rue, etc. ? Les résultats seront fonction des définitions que l'on aura établies au point de départ.

3.5 Représentation graphique

Concluons sur ce point en parlant de la représentation graphique qui est également différente selon la discipline. En mathématique aussi bien qu'en statistique, l'ordinateur a apporté une plus grande utilisation des représentations qui, dans le passé, étaient soit trop longues à faire ou presque impossibles à réaliser. Aujourd'hui, l'ordinateur permet par exemple de visualiser facilement des fonctions paramétriques ou des animations à trois dimensions. Toutefois, en analyse de données, même avec des outils de représentation très simples comme les diagrammes « tiges et feuilles » et les « boîtes à moustaches », on peut obtenir différentes informations en se servant de plusieurs représentations (Biehler, 2008 ; Shaughnessy, 2006). L'exemple suivant tiré de Cobb et Moore (2000, p. 807-808) l'illustre bien. Il provient de l'analyse exploratoire de données d'une expérimentation où l'on étudiait l'impact de la prise de calcium sur la réduction de la pression sanguine systolique par deux groupes d'hommes (placebo et calcium).

Placebo		Calcium
1	-1	
5	-0	5
33111	-0	234
3	0	1
5	0	7
2	1	01
	1	78

FIGURE 1[4] – *Diagrammes « tiges et feuilles » en parallèle de la réduction de la pression systolique du sang chez deux groupes d'hommes (Cobb & Moore, 2000, p. 808)*

Dans la représentation en parallèle de deux diagrammes « tiges et feuilles » (Figure 1), on remarque que dans le groupe prenant du calcium, il y a une apparence de bimodalité qui pourrait indiquer la présence de deux sortes de sujets. Pourrait-il y avoir deux types de réponses à la prise de calcium ? La possibilité est à explorer. Si l'on représente ces mêmes données avec des boîtes à moustaches (Figure 2), l'examen de ces représentations révèle une différence importante dans les médianes mais aussi dans la dispersion des données, ce qui sème des doutes sur l'hypothèse de variances égales pour la suite de l'analyse. Ces points sont facilement observables et ces analyses exploratoires deviendront très utiles quand viendra le temps d'utiliser des tests.

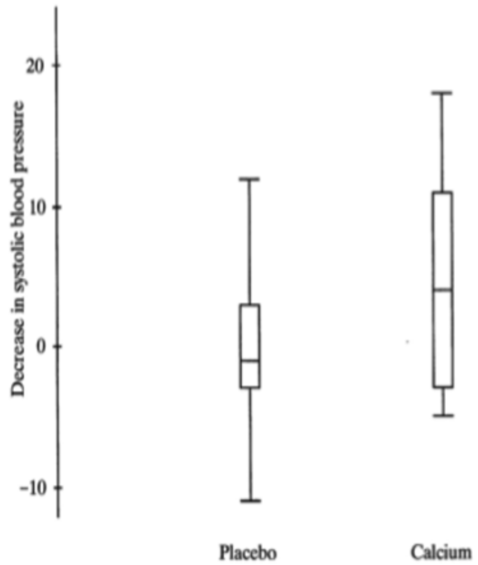

FIGURE 2[5] – Boîtes à moustaches en parallèle de la réduction de la pression systolique du sang chez deux groupes d'hommes (Cobb & Moore, 2000, p. 809)

[4] Le diagramme « tiges et feuilles », outil de l'analyse exploratoire des données statistiques, a en plus une valeur didactique pour la mathématique. Sa construction peut permettre, entre autres, de consolider le concept de nombre (Gattuso, 2006b), ce qui n'est pas à négliger dans un contexte où la statistique est enseignée dans le cours de mathématique.

[5] La boîte à moustaches résume seulement quelques caractéristiques de position du caractère étudié

On pourrait ajouter la représentation de la droite de Henry qui est une méthode graphique pour ajuster une distribution gaussienne à celle d'une série d'observations (d'une variable numérique continue). En cas d'ajustement, elle permet de lire rapidement la moyenne et l'écart-type d'une telle distribution et est relativement facile à réaliser si on fournit le papier gausso-arithmétique. N'oublions pas les diagrammes de dispersion qui, dans le cas d'études à deux variables, sont indispensables et peuvent apporter des informations utiles avant même d'appliquer les calculs de corrélation et d'ajustement de droite.

4 Quand ?

Il n'y a pas encore très longtemps, on repoussait l'enseignement de la statistique au niveau universitaire. Bien que cela ait changé, certains professeurs prétendent que vaudrait mieux « pas du tout, que mal! » en rapportant que leurs étudiants arrivent trop souvent avec des connaissances et des conceptions erronées. Cependant, il est bien connu, par exemple, que l'apprentissage d'une langue est plus facile s'il est commencé dès l'enfance ; pourquoi cela ne serait-il pas vrai de la statistique ?

Girard (2004) fait un parallèle entre la construction du modèle euclidien de géométrie et l'apprentissage des probabilités. Alors que, pour le modèle euclidien, on prévoit un apprentissage graduel échelonné sur une dizaine d'années, pour les probabilités, le processus d'apprentissage s'étend sur un ou deux ans. En effet, depuis le premier âge, l'enfant joue avec des cubes qu'il aura la possibilité de manipuler abondamment. Dès le début de la scolarité, il aura l'occasion de représenter cet objet par des dessins, puis graduellement, on passera aux propriétés et éventuellement à l'objet mathématique.

Alors pourquoi n'offrirait-on pas le même type de parcours pour les probabilités et, pourquoi pas, pour la statistique comme suggéré par l'auteur ? Surtout qu'une présentation tardive qui arrive « quand de nombreuses conceptions plus ou moins adéquates, voire erronées, se sont installées chez les élèves » (Girard, *id.*, p. 87) n'arrange pas les choses.

Mais quand alors doit-on commencer et est-ce vraiment possible ? Avec Scheaffer (2002), je crois que cela doit commencer à l'école et faire partie de la formation à tous les niveaux. Des expériences récentes montrent qu'on peut le faire avec succès même à la maternelle. Il faut évidemment à l'enseignant une préparation qui permette d'exploiter les productions des élèves et de les guider dans leur apprentissage.

Schwartz & Whitin (2006) relatent une expérience menée avec des enfants de 4 ans dans une maternelle. On offrait aux enfants l'occasion d'en savoir plus sur leurs petits camarades sur un sujet de leur choix. Voici le cas d'Harold : Harold voulait savoir qui, dans la classe, avait appris à lacer ses chaussures. La maîtresse ayant approuvé cette question lui donne une feuille blanche pour réaliser cette tâche. « Blanche » pour laisser à l'enfant toute la latitude possible pour enregistrer les informations qu'il recueillera. Voici le résultat (Figure 3).

(médiane, quartiles, minimum, maximum ou déciles). Ce diagramme est utilisé principalement pour comparer un même caractère dans deux populations de tailles différentes. Il s'agit de tracer un rectangle allant du premier quartile au troisième quartile et coupé par la médiane. Ce rectangle suffit pour le diagramme en boîte. On ajoute alors des segments aux extrémités menant jusqu'aux valeurs extrêmes, ou jusqu'aux premier et neuvième déciles, voire aux 5^{e} et 95^{e} centiles. On parle alors de diagramme en boîte à moustaches ou de diagramme à pattes. (http://fr.wikipedia.org/wiki/Boîte_à_moustaches).

De prime abord, ceci ressemble fort à des gribouillis ! Mais l'enseignante bien préparée a demandé à Harold de lui expliquer ses résultats. Le garçon avait d'abord décidé de demander à chacun de ses compagnons de lui montrer qu'ils étaient capables de lacer leurs chaussures en le faisant devant lui. Il explique qu'après réflexion, il a découpé en deux la procédure.

FIGURE 3 – *Le graphe de Harold. (Schwartz & Whitin, 2006, p. 8)*

Il a utilisé le H, la première lettre de son nom, la seule qu'il connaissait, pour noter ceux qui savaient lacer leurs chaussures. Il a utilisé des symboles en boucles c, ℓ (voir le coin supérieur gauche de l'illustration) pour ceux qui pouvaient aussi nouer les lacets pour faire la boucle et × pour ceux qui ne savaient pas. Une non-réponse a été notée ○. L'enfant a : 1) posé sa propre question, 2) démontré une correspondance un-à-un, 3) distingué plusieurs niveaux d'informations, 4) inventé des symboles, 5) résolu un problème d'enregistrement de données, 6) vérifié la validité des réponses et 7) organisé (bien qu'élémentairement : 4 ans !) l'enregistrement des données. Cependant, cela n'est devenu évident que lorsque l'enseignante lui a demandé d'expliquer ses résultats. Il fallait de plus qu'elle soit à même de voir toute la richesse de ceux-ci.

Une autre enfant a voulu savoir quelle activité préféraient ses amis (Figure 4). L'enfant explique qu'elle a noté ses résultats par des dessins représentant chacune des activités. Elle a représenté horizontalement ses données montrant qu'elle avait déjà à l'esprit un ordre de catégories. Elle a noté le nombre de fois que chaque activité a été choisie (l'effectif) et elle a pu relier chacune des réponses à un des enfants de la classe.

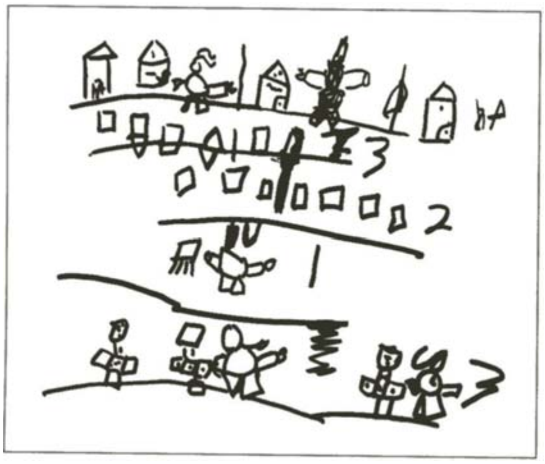

FIGURE 4 – *Activité préférée. (Schwartz & Whitin, 2006, p. 12)*

Un troisième a demandé quel était leur métier préféré (Figure 5). Il a représenté chaque métier par un dessin de couleur différente : pompiers en rouge, policiers en bleu, boulangers en orangé, enseignants en brun. Il a fait un grand pompier au centre pour montrer que c'était le métier le plus souvent choisi (le mode) et a enregistré à qui la demande avait déjà été faite en demandant à chacun leur carte de visite[6], évitant ainsi de reposer deux fois la question à un même sujet.

FIGURE 5 – *Métier préféré. (Schwartz & Whitin, 2006, p. 13)*

5 Comment ?

5.1 Le parcours du statisticien

On ne parle plus d'enseigner seulement des concepts mais de développer une pensée et un raisonnement statistique. Il faut aller au-delà des calculs de la moyenne ou même de l'intervalle de confiance qui sont trop souvent présentés comme des formules à appliquer. Bien qu'elle fasse grand usage de mathématiques, la statistique devrait être enseignée comme telle et non comme la mathématique (Moore & Cobb, 2000). Les élèves doivent développer une compréhension conceptuelle. Savoir effectuer une série de calculs n'assure pas la compréhension des concepts statistiques, des formules ou de leurs applications dans la réalité.

Pour cela, les enseignants doivent aider leurs élèves à :

1. donner un sens à l'univers quotidien en utilisant les données ;
2. comprendre la variabilité et quantifier l'incertain ;
3. planifier et concevoir des expérimentations (ou enquêtes) de façon à s'assurer que les données recueillies soient utiles pour répondre à la question posée ;
4. tirer des conclusions à partir d'échantillons considérant la manière dont les données ont été recueillies ainsi que leur incertitude et leurs sources de variabilité ;
5. utiliser différents outils technologiques pour permettre des explorations plus réalistes des données. (Burrill & Elliott, 2006).

[6] Les enfants de la classe avaient des petites cartes de visite où était inscrit leur nom.

Les travaux de Tukey (1977), entre autres, et l'accessibilité d'outils technologiques ont rendu possible l'expérimentation du parcours du statisticien même à l'école. C'est par ces expérimentations que la pensée statistique se construira. Il faut absolument travailler avec des données et éviter au point de départ de faire trop de théorie, la plupart des calculs pouvant être aujourd'hui automatisés. À la base, il faut s'en tenir à l'analyse exploratoire des données qui va plus loin que la statistique descriptive et qui, de toute façon, est préliminaire à l'inférence. La technologie a donné des outils avancés utilisables par tous, mais soyons prudent car ils peuvent aussi être utilisés par ceux qui ont peu de connaissance de la statistique ou même de la mathématique, ce qui rend la compréhension encore plus nécessaire.

Il faut insister sur les grandes idées statistiques, la variabilité (son observation, sa mesure) et la distribution qui sont au cœur de la statistique. Il reste que certains concepts, et en particulier ceux qui font appel au raisonnement proportionnel, sont plus difficiles, sans oublier que les élèves construisent leurs connaissances à partir de leurs expériences antérieures et de leurs conceptions qui peuvent être erronées (*misconceptions*).

Tout un programme ! Faire tout cela en se rapprochant le plus possible du processus relié à la statistique parcouru par le statisticien professionnel :

- questionnement ;
- planification de l'expérience (expérimentation ou enquête[7]) ;
- collecte de données ;
- compilation/organisation des données et représentation ;
- analyse, interprétation (éventuellement, inférence) ;
- conclusion et communication des résultats ;
- re-questionnement : esprit critique.

5.2 La question

La définition et la formulation de la question est à la base de l'expérimentation (ou de l'enquête). La suite en dépend. Il faut apprendre à formuler la question de sorte à trouver des réponses à ce que l'on veut savoir.

L'enseignant doit laisser les élèves libres de poser les questions qui les intéressent et de trouver le moyen d'y répondre, tout en les questionnant pour les amener à clarifier la formulation de leur(s) question(s) sans oublier la question originale et ce qui les intéresse. Il doit les amener à examiner la méthode de collecte pour voir si les données obtenues apporteront une réponse à la question initiale. Il peut leur suggérer de tester les questions[8] et de les raffiner s'il y a lieu (Russell, 2006).

Par exemple : « Combien y a-t-il de maisons dans ta rue ? ». La question est-elle comprise de tous ? Pour que les données recueillies soient valides, il faut aussi que la question soit précise. « Comment compte-t-on les maisons ? » Faut-il comptabiliser aussi celles qui sont en construction ? Les appartements et les copropriétés également ? Il faut se demander exactement qu'est-ce que l'on veut savoir et s'assurer que les réponses à la question apporteront l'information adéquate. Les données obtenues doivent être significatives.

[7] Les exemples cités sont plus souvent des enquêtes (plus faciles à réaliser à l'école) mais pourraient tout aussi bien être des expérimentations donnant lieu à une collecte de données.

[8] Il peut y avoir une ou des questions, et il peut aussi s'agir d'une expérimentation plutôt qu'une enquête. Les principes demeurent les mêmes.

London-McNab (2006) relate une expérience où cette étape a été très importante. Les élèves voulaient ordonner les pays du Commonwealth d'après leurs résultats aux Jeux du Commonwealth[9]. Devait-on compter le nombre de médailles d'or ? Ou encore l'ensemble des médailles ? Ne devrait-on pas tenir compte de la population du pays ? Ou encore du PIB ? Les pays plus riches n'avaient-ils pas au point de départ une avance ? Que dire du climat ? Les pays de climat tropical étaient probablement désavantagés dans une épreuve « nordique ». Les élèves ont donc dû définir des critères et développer un codage pertinent et objectif pour ordonner les pays participant aux Jeux. Les élèves ont conclu en communiquant leurs résultats au Comité organisateur des Jeux qui leur a répondu. Un bel encouragement pour leur travail !

5.3 La planification

Une fois la question définie, il faut planifier la suite. À qui posera-t-on la question ? Il faudra éventuellement distinguer la population de l'échantillon et parler des techniques d'échantillonnage. Il est aussi très intéressant de pouvoir procéder à des comparaisons, donc d'utiliser plusieurs échantillons. Par exemple, si on pose la question « Etes-vous en faveur du port obligatoire d'un uniforme à l'école ? » aux enfants et à leurs parents, le poids des parents pourrait être plus fort (si les enfants ont deux parents). Alors, dans ce cas, on peut séparer les réponses des parents de celles des enfants et comparer les deux groupes.

Comment enregistrer les résultats ? Nous avons vu que l'utilisation d'une page blanche avec les enfants de maternelle leur laissait plus de latitude, mais quel que soit l'outil, l'ordinateur, le papier, la photo ou autre, il faut le décider avant de commencer.

5.4 La collecte des données

Comment cueillir les données ? La question sera-t-elle ouverte ? Ou à réponses multiples ? Si on demande aux enfants leur collation préférée et que l'on ne met que des fruits comme choix de réponse, peut-être passera-t-on à côté de leur collation favorite.

Les données recueillies sont essentielles puisque la statistique est l'étude des données, mais il faut comprendre qu'elles ne sont qu'une partie de la réalité. Il faut savoir ce qu'elles veulent dire pour ne pas les interpréter au-delà de leur domaine. Au passage, l'enseignant soulignera le type des données : quantitatives, ordinales, qualitatives. On examinera comment elles ont été mesurées, dans quelle condition la collecte a été faite, comment l'échantillon a été choisi afin de pouvoir relativiser les conclusions au bout de l'analyse. Une expérimentation mal planifiée peut invalider les résultats (Teague, 2006).

5.5 La représentation des données

Pour pouvoir analyser les données, il est essentiel de les assembler et de les organiser. Nous avons vu que même les enfants très jeunes peuvent arranger les données cueillies de sorte à pouvoir en tirer des informations. Mais lorsqu'elles sont plus nombreuses, il faut être attentif au moyen de les ordonner et de les classer. Il faut aussi pouvoir retourner aux données brutes pour vérification ou encore pour voir si une donnée « aberrante » est due à une erreur de transcription ou pas.

[9] Les jeux du Commonwealth sont comme les Jeux Olympiques mais limités aux nations faisant partie du Commonwealth.

Éventuellement, les données seront organisées en tableaux. Alors que cela semble si simple, quand on arrive aux données numériques, les erreurs sont fréquentes, les enfants voient des nombres dans le tableau et oublient leur signification. Dans des recherches sur la moyenne, nous avons souvent vu des élèves additionner les valeurs de la variable ou encore les effectifs pour en faire la moyenne (Gattuso & Mary, 2001, 2005a, 2005b). C'est pourquoi il est essentiel tout au long du parcours de rester près du sens des données, de connecter ces nombres à la réalité en faisant verbaliser les enfants.

TABLEAU 1 – *Nombre d'enfants dans la famille des enfants de la classe*

Nb d'enfants dans la famille	Effectif
1	8
2	4
3	6
4	1
5	0
6	1

« Combien d'enfants de la classe viennent de familles de 3 enfants ? », « 1 enfant de la classe a 4 enfants dans sa famille. », etc. Ceci est très important pour que l'enfant saisisse la différence entre, par exemple, le 4 qui veut dire : « 4 enfants de la classe viennent d'une famille de 2 enfants» du 4 qui veut dire : « Il y a 1 enfant dans la classe qui est dans une famille de 4 enfants»[10]. Il faudra encore insister davantage quand les effectifs relatifs, les pourcentages, seront utilisés. Cette verbalisation ne doit pas être réservée seulement aux petits[11].

Les tableaux ne sont qu'une façon d'organiser et de présenter les données. Les diverses représentations graphiques sont tout aussi essentielles. Dès le plus jeune âge, des expériences le montrent (Perelli-D'argenzio *et al.*, 2006), les enfants laissés à eux-mêmes construisent leurs propres représentations qui, bien que non conventionnelles, ont du sens. Ce n'est que lentement qu'on devrait les amener à lire d'abord et adopter ensuite des représentations plus conventionnelles, telles les diagrammes « tiges et feuilles » (Dunkell, 1986), les boîtes à moustaches, sans parler des classiques diagrammes en bâtons, tartes (ou camemberts), histogrammes, etc. Alors que l'on considère trop souvent ces représentations comme faciles, elles comportent plusieurs difficultés (Jelinski, 1993). Déjà dans la boîte à moustaches, on a les quartiles et la médiane de la distribution et les valeurs extrêmes, minimum et maximum. Les diagrammes circulaires sont basés sur la proportionnalité et l'histogramme demande un passage du linéaire à l'aire particulièrement lorsque les classes ne sont pas toutes de largeur égale (ce que l'on évite trop souvent à l'école). C'est un passage important pour la compréhension éventuelle des graphes des fonctions de densité, comme la « cloche » de Gauss.

Non seulement on devrait être capable de représenter les données mais aussi de juger laquelle ou lesquelles de ces représentations informent le mieux car, en statistique, une représentation peut faire ressortir des informations qui sont masquées dans une autre, comme nous l'avons déjà souligné.

[10] Nous ne reviendrons pas ici sur l'importance de clarifier la question. Dans les familles recomposées d'aujourd'hui, il y aurait lieu de préciser qui est « enfant » de la famille.

[11] Ces erreurs ont été observées chez des élèves plus âgés (15 ans et plus) dans les recherches sur la moyenne (Gattuso & Mary, 2001, 2005a, 2005b).

5.6 L'analyse

Ayant organisé les données, il faut les « regarder » et, surtout, les regarder comme un tout, comme une distribution. Il faut en examiner la forme et la décrire : y-a-t-il des bosses ou une bosse ? Y a-t-il des données « aberrantes » ? (Bakker, 2001). Mettre en parallèle et comparer les résultats de deux groupes est fortement recommandé pour aider à distinguer les caractéristiques et à faire ressortir les éléments importants des distributions. L'enseignant doit questionner les élèves pour les amener à considérer la distribution comme un tout. Il doit les encourager à dire où les données sont concentrées. Par exemple, au lieu de leur faire des demandes très ponctuelles : « Combien de fois a-t-on obtenu 6 en lançant nos deux dés ? », on peut demander : « Où se trouve la masse des résultats ? Entre quelles valeurs de la somme est-on le plus souvent tombé en lançant les deux dés ? » (Russell, 2006).

Le choix des concepts enseignés est aussi important. Il y a d'autres mesures de tendance centrale que la moyenne, tels la médiane, le mode et, pour la dispersion, l'étendue, la distance interquartile plus accessibles que l'écart-type. D'ailleurs dans des cas simples, ces mesures, mode, médiane, quartiles, peuvent être trouvées sans l'utilisation de formules à appliquer mais en donnant un sens à ce que l'on fait. Par exemple, pour trouver une droite qui traduit le mieux possible la relation entre deux ensembles de données, il n'y a pas que la méthode formelle des moindres carrés, mais aussi la méthode des médianes ou l'approche par moyennes locales, sans oublier que la somme minimale des carrés des écarts entre les points du nuage et ceux de la droite peut être trouvée en jouant sur différents paramètres en utilisant un chiffrier (Gattuso, 2008a)[12].

Tout au long du parcours scolaire, il serait bon de préparer les élèves à l'inférence sans pour cela avoir recours à des outils compliqués. Dans une classe, des élèves en difficulté sont arrivés avec la question suivante : « Est-ce que les gens qui portent des lunettes regardent plus la télévision que les autres ? » (Mary & Theis, 2007). Ils ont ensuite décidé de séparer ceux qui regardaient 20 heures ou moins de télé par semaine de ceux qui la regardaient plus de 20 heures. Un groupe a même produit un tableau à deux dimensions. Les résultats de cette classe sont les suivants :

TABLEAU 2 – *Heures de télévision regardées par les enfants portant des lunettes ou pas*

Heures de télé/sem.	Lunettes	Pas de lunettes	Total
0-20	4	10	14
20+	6	5	11
Total	10	15	25

À partir de ce tableau, on peut demander aux élèves de faire un diagramme en bâtons pour comparer les résultats de ceux qui portent des lunettes ou pas (Figure 6). On remarque alors que la situation de ceux qui portent des lunettes est différente de celle de ceux qui n'en portent pas. On voit clairement que les enfants de la classe qui ne portent pas de lunettes sont plus nombreux à regarder moins la télévision (sans oublier dans ce cas qu'il y en a presque

[12] Au point de départ, c'est l'idée qui compte. À la limite, on peut initier les élèves en leur faisant jouer avec une baguette sur leur feuille, la déplaçant pour approcher le plus possible les points du nuage, pour ensuite tracer une droite et possiblement en trouver l'équation. On cherchera ensuite qui a la plus petite somme des carrés des écarts. Ils seront ensuite plus motivés à connaître des méthodes peut-être plus complexes mais aussi plus efficaces.

autant dans les deux groupes – lunettes ou pas – qui regardent beaucoup la télé). On peut dire que, « selon les chiffres », la situation entre les deux groupes est différente, mais est-ce que cela dépend du port de lunettes ou pas ? On peut amorcer la discussion sans aller plus loin que ce que les élèves sont à même de comprendre. Pour des élèves plus avancés, on peut aussi passer aux effectifs relatifs (quel dénominateur ?) et aux pourcentages (pourcentage de quoi ?). Les résultats seront-ils les mêmes ? Ceci suppose que l'enseignant ait les connaissances pour mener à bien la discussion car cet exercice veut préparer la voie aux tableaux de contingence (ou à double entrée) et à la chi-carrée qui sera travaillée dans les cours plus avancés.

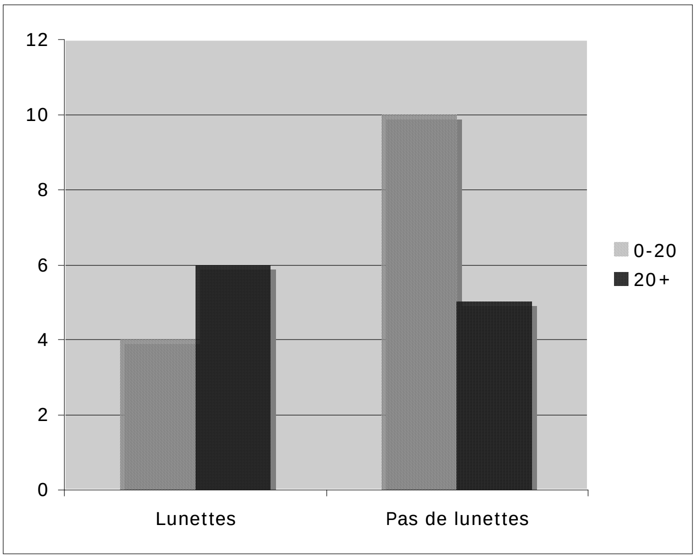

FIGURE 6 – *Heures de télévision regardées par les enfants portant des lunettes ou pas*

Tarr *et al*. (2006) relatent une activité qui s'est déroulée dans une classe de 6e année (11 ans). La situation présentée demandait de tester les dés provenant de 6 compagnies différentes pour en évaluer la qualité. Les élèves devaient collecter des données pour voir si les dés étaient équilibrés ou non, autrement dit pour vérifier si les 6 résultats possibles étaient équiprobables. On leur présentait les questions suivantes :

– Les dés sont-ils équilibrés, en recommandez-vous l'achat ?
– Quelle validation convaincante (*compelling evidence*) avez-vous pour dire que les dés que vous avez testés sont équilibrés ou pas ?
– Utilisez vos données pour estimer la probabilité de chaque résultat, de 1 à 6, pour les dés testés.

On leur demandait aussi de produire des graphiques, un poster et de faire ensuite la présentation de leurs résultats.

Plusieurs questions ont surgi pendant l'activité. Comment tester les dés ? En les lançant ? Mais combien de fois ? Quelle est la taille d'échantillon acceptable ? Petit ou grand échantillon ? Est-ce légitime de mettre ensemble plusieurs essais ? Ensuite, quelle variation peut-on tolérer dans les résultats ? Peut-on utiliser des simulations par ordinateur ? Qu'est-ce qu'une validation convaincante ? Si on répète l'expérimentation, doit-on s'attendre à de la variabilité ?

À ce niveau, le rôle de l'enseignant n'est pas de fournir des réponses « finales » mais bien d'écouter les élèves, de les questionner et de favoriser la discussion en classe pour amener les élèves à défendre leurs positions et éventuellement à les remettre en question. En fin de parcours, les élèves de l'expérimentation ont proposé des réponses aux questions de départ tout en étant conscients de la relativité de ces résultats.

Toutefois, pendant cette activité, les élèves ont pu :

- formuler la question à laquelle on pouvait répondre par des données. La demande était « Quel dé est équilibré ? » et ils l'ont traduite par la question : « Les résultats de 1 à 6 ont-ils la même chance de se produire ? », question à laquelle ils pouvaient répondre par une expérimentation ;
- collecter des données et raisonner à partir de celles-ci ;
- organiser et présenter les données appropriées pour répondre à la question ;
- choisir et utiliser les méthodes statistiques appropriées pour analyser les données ;
- développer et évaluer des « inférences » et des prédictions basées sur les données ;
- présenter les données et communiquer leurs conclusions et leurs arguments.

Plus précisément, ils ont pris conscience de l'influence de la taille de l'échantillon, ce qui est souvent omis dans les programmes scolaires. Ils ont pu constater que les différences entre la prédiction et les résultats s'ils sont tirés d'un grand échantillon représentatif doivent être prises au sérieux.

Ces raisonnements sèment un germe et préparent à long terme l'apprentissage des notions de distribution échantillonnale, d'intervalle de confiance, de test, de loi des grands nombres, favorisant ainsi le développement du raisonnement et de la pensée statistique.

5.7 La conclusion et la communication

Quelle que soit l'importance du travail, il faut, en fin de parcours, favoriser la communication (orale ou écrite) de conclusions appuyées sur des arguments quantitatifs reliés à des données et à un contexte réel et illustrés de graphiques ou/et de tableaux, car un nombre en soi n'a pas de sens en analyse de données. Selon Groth (2006), les élèves impliqués dans des analyses de données authentiques seront plus réticents à conclure de façon trop simpliste. La tâche des enseignants, quel que soit le niveau, est toujours de poser des questions, de mettre au défi les élèves, de les amener à utiliser des justifications basées sur les données et sur le contexte plutôt que des solutions trouvées en appliquant aveuglément des procédures de calculs statistiques.

5.8 De nouvelles questions : esprit critique

Encore une fois, il faut regarder avec esprit critique les résultats, mais aussi la question, la collecte, les données, en somme, faire un retour en arrière sans oublier qu'il est essentiel d'interpréter tous ces résultats dans le contexte, sans quoi ils n'ont pas de sens. Ceci peut permettre d'accepter qu'il peut y avoir des conclusions différentes selon les options priviligiées au cours du parcours. Analysant les résultats, il faut favoriser la discussion, le retour critique et accepter de mettre en doute les choix qui ont été faits. « Notre question aurait-elle pu être précisée, ou différente ? A-t-on choisi la bonne population (ou échantillon) ? A-t-on fait un assez grand nombre d'essais ? ». Ainsi de suite.

5.9 Évaluation

Un dernier point, l'évaluation. En effet, on constate fréquemment que l'évaluation dicte l'enseignement. Et c'est pourquoi il est primordial d'y être attentif. Trop souvent les questions posées en exercice ou en contrôle sont plus mathématiques que statistiques. Par exemple, « le poids moyen de 50 tomates est 2,36 livres, quel est le poids total ? ». Il est vrai qu'il faut connaître la relation entre le total et la moyenne, mais il n'y a pas d'analyse de données et le contexte est loin des préoccupations des élèves. Mais plus encore, citons un item où l'on raconte que des élèves ont choisi parmi trois livres celui qu'ils préféraient. Les indices donnés sont ceux-ci : a) il y avait 34 élèves ;b) le gagnant a eu le plus grand nombre de votes mais moins que la moitié ; c) les deuxièmes choix étaient ex-aequo. La question posée était : « quels sont les résultats du sondage ? » Ici, on se pose des questions au sujet de données, mais cela ne requiert pas un raisonnement statistique mais bien un raisonnement sur la propriété des nombres.

Un exemple plus adéquat, tiré de Scheaffer (2006), pose une question ouverte et utilise des données réelles (ou du moins réalistes) pour répondre à une question pratique. Elle demande un raisonnement statistique puisqu'il y a plusieurs réponses plausibles dont aucune n'est la « meilleure » à tous les points de vue. De plus, les mathématiques nécessaires sont élémentaires.

« Le tableau suivant donne les temps de courses de 100 mètres pour chaque fille. On doit en choisir une seule pour la prochaine compétition. Laquelle choisiriez-vous et pourquoi ? (Burrill & Collins, 2005, p. 201-204; voir Scheaffer, 2006)

TABLEAU 3 – *Temps de course pour trois coureuses de cent mètres*

Course #	1	2	3	4	5	6	7
Louise	15,2	14,8	15,0	14,7	14,3	14,5	14,5
Noémie	15,8	15,7	15,4	15,0	14,8	14,6	14,5
Martine	15,6	15,5	14,8	15,1	14,5	14,7	14,5

Dans un autre exemple, toujours tiré de Scheaffer (2006), on présente un tableau qui fournit les données (pour 10 ans) de la population porcine (en milliers) et l'ammoniaque atmosphérique (en parties pour million) mesurée dans des fermes porcines par le Programme National de Dépôt Atmosphérique Américain. On demande d'abord de construire un diagramme de dispersion. On demande d'interpréter le coefficient de corrélation qui est donné : 0,85. En s'appuyant sur la valeur de ce coefficient et sur le diagramme, on demande si la relation semble linéaire et pourquoi. Finalement, on demande si l'on peut dire que l'ammoniaque atmosphérique dans les fermes peut être expliqué par la taille de la population porcine. Dans ce problème, le contexte est clair, les données sont réelles et répondre demande peu de calculs mais un raisonnement statistique qui requiert une compréhension de la différence entre association et causalité (qui ne peut pas être déterminée avec ces seules données).

5.10 En conclusion

Une même tâche peut être adaptée à divers niveaux de difficulté, selon l'âge des élèves et en tenant compte du temps alloué. Il est important qu'elle leur soit accessible mais non plus trop facile ; il faut susciter leur intérêt. Les enquêtes ne doivent pas toujours être compliquées et élaborées, on peut faire une activité rapide de quinze minutes mettant en place la plupart des étapes indiquées précédemment.

Cependant, il est important de ne pas se limiter à un seul concept : « Cette année, on fait la moyenne! … ». Il est préférable d'intégrer les concepts visés dans une activité plus large et d'avoir en vue ce qui viendra plus tard, comme l'inférence par exemple. Il ne faut pas avoir peur d'utiliser du matériel concret sans négliger l'ordinateur qui est un support de plus en plus indispensable pour le développement du raisonnement statistique.

La tâche proposée devrait également permettre de soulever les conceptions justes ou inadéquates des élèves et donner l'occasion aux enseignants d'avoir accès au raisonnement des élèves. Les enseignants devraient tailler sur mesure leur questionnement et se préparer à poser des obstacles et à guider les élèves vers un raisonnement de niveau supérieur.

Et surtout, ne pas oublier que les élèves sont souvent plus forts qu'on ne le croit !

Tout cela demande beaucoup des enseignants qui sont en première ligne et qui généralement ont peu d'expérience et peu de formation pour faire face à ce type de tâche.

6 Pourquoi pas ?

6.1 Pourquoi cela ne se fait pas ?

Lors d'une expérimentation dans des écoles italiennes[13], des entrevues, faites auprès d'enseignants de mathématique au lycée, révèlent « une sensation d'insécurité due à un manque de préparation en statistique, mais aussi en didactique de la statistique (Gattuso & Pannonne, 2000, 2002). Par exemple, au Québec où la formation des enseignants est de 4 années universitaires, peu ou pas de temps est consacré spécifiquement à la statistique ou à sa didactique. Au niveau secondaire, les futurs enseignants suivront un cours d'un semestre en statistique souvent plutôt théorique et difficilement transférable dans un enseignement tel que souhaité.

Par conséquent, les enseignants ne se sentent pas à l'aise avec la statistique. Même s'ils ont suivi un cours de statistique, ils ne la travaillent pas en classe par manque de confiance en soi. L'obligation d'enseigner un sujet pas très bien maîtrisé est pour eux génératrice d'anxiété.

N'ayant pas de formation spécifique, les enseignants doivent se fier aux manuels scolaires sans pouvoir y déceler les erreurs présentes (Hawkins *et al.*, 1992). Or, les manuels examinés n'apportent pas le support nécessaire aux enseignants (Queiroz & Coutinho, 2006). De plus, le manque de formation des enseignants en didactique de la statistique font qu'ils sont incapables en classe de reconnaître et de répondre aux conceptions justes ou inadéquates de leurs élèves (Russell, 1990).

D'autre part, les enseignants ne sont pas toujours conscients de la richesse des contenus statistiques qu'ils doivent enseigner. La statistique, et particulièrement les statistiques descriptives (faciles et sans intérêt !), sont vues comme un ajout sans importance. Conséquemment, la statistique est reléguée à la fin de l'année scolaire sinon complètement

[13] 2 130 élèves et 189 enseignants du primaire, 1 632 étudiants et 86 enseignants du secondaire et 2 314 étudiants et 107 enseignants de l'école supérieure ont participé et complété toutes les étapes de l'expérimentation.

évacuée par manque de temps (Aksu, 1990). Ce dernier argument, selon nous, perd sa force si l'on souligne toute la mathématique travaillée dans le cadre de la statistique (Gattuso, 2008a).

6.2 Pourquoi on ne le ferait pas ?

6.2.1 Formation des enseignants

La nécessité de la statistique pour la formation du citoyen moyen est dorénavant reconnue presque partout et son enseignement a été intégré aux programmes scolaires à différents degrés. Il faut maintenant s'attaquer à la formation des enseignants qui est actuellement le point faible de la chaîne. C'est la clé du développement de l'enseignement de la statistique. Les enseignants sont le lien entre le savoir et l'élève ; sans une intervention adéquate de leur part, la pensée et le raisonnement statistiques risquent de rester inconnus à leurs élèves.

Le manque de connaissance en statistique des enseignants est largement documenté. Les enseignants ne sont pas familiers avec l'analyse des données (Russell, 1990) mais aussi avec des concepts élémentaires comme la moyenne et la médiane (Estrada & Batanero, 2008 ; Jacobbe, 2008). À ces difficultés, s'ajoutent des problèmes avec les graphes, avec les concepts de variation et de distribution (Espinel *et al*., 2008; Canada, 2008). La statistique est l'étude des données et la plupart des enseignants du primaire ont peu d'expérience dans ce champ et ce, malgré la présence de données dans les medias. Les futurs enseignants et les enseignants ont besoin de connaissances techniques et conceptuelles de la statistique et de la probabilité qui sont maintenant au programme au primaire. (Scheaffer, 2002). Ils ont aussi peu de pratique de la simulation et de l'expérimentation, ils connaissent peu les graphes, pas du tout les diagrammes « tiges et feuilles » et les boîtes à moustaches, et rien des difficultés d'apprentissage qui y sont reliées. La plupart n'ont pas eu l'occasion de développer leurs propres pensée et raisonnement statistique (Rossman *et al.*, 2006). En résumé, les enseignants ont peu ou pas de connaissance de la statistique et presque jamais de formation en enseignement de la statistique (didactique) (Batanero, 2010). De plus, selon Estrada et Batanero (*id.*), l'attitude négative des enseignants face à la statistique s'accroît avec l'expérience d'enseignement mais, (heureusement) elle s'améliore avec l'augmentation de la connaissance.

6.2.2 Connaissances didactiques

Personne ne discutera la nécessité d'avoir une bonne connaissance et une bonne compréhension de la statistique pour pouvoir l'enseigner. Cependant, ce n'est pas suffisant de connaître la théorie, il faut ajouter à cela des connaissances didactiques ou ce que les Américains appellent le « pedagogical content knowledge » (Shulman,1986).

Les enseignants doivent comprendre comment les élèves apprennent les concepts statistiques (conceptions, difficultés, raisonnements : didactique). Pour pouvoir aider les élèves, ils ont besoin de connaître les erreurs communes, les conceptions erronées, les stratégies les plus souvent utilisées par les apprenants et les différents stages de développement des concepts.

Il leur faut aussi avoir un bagage d'illustrations et savoir les utiliser pour amener les élèves à développer leur sens critique. Des exemples improvisés ne fonctionnent pas toujours. Les décisions prises dans la classe dépendent des connaissances des enseignants de la didactique de la statistique et de la statistique elle-même, ainsi que de leur capacité à jongler avec l'ambiguïté et l'incertitude caractéristiques de la démarche statistique. Une connaissance des erreurs communes ou des différentes stratégies utilisées aiderait les enseignants à préparer des activités qui confrontent ces erreurs.

Une formation adéquate permettrait aux enseignants de construire ou d'adapter des tâches pour préparer leurs leçons. Pour cela, ils doivent être à même d'examiner de façon critique des ensembles de données, de reconnaître les concepts qui peuvent être travaillés avec un certain ensemble de données et ainsi mettre en œuvre un apprentissage efficace.

Ils doivent demander à leurs élèves des explications et pas seulement une réponse qui parfois cache une justification complètement fausse. Par exemple, dans une expérimentation, on a demandé à 173 enfants de 11 ans et plus de répondre à la question suivante : « Si un garçon lançait deux pièces de monnaie, quelle est la probabilité qu'il obtienne une pile et une face ? Expliquez. » Alors que plus de la moitié (93/173) ont répondu ½, 40 d'entre eux avaient des justifications erronées (Rubel, 2006).

Il y a également un besoin d'activités significatives qui permettent aux élèves de travailler avec des enquêtes statistiques. Or, les manuels utilisent rarement des données réelles.

Comment préparer les enseignants à cette tâche ? Plusieurs voies ont été mises en place avec succès. Lors d'une expérimentation à l'échelle nationale en Italie (CIRDIS, 1998-2000), la formation des enseignants aux différents niveaux (élémentaire, secondaire et lycée) s'est faite en faisant expérimenter aux enseignants ce que l'on voulait faire expérimenter aux élèves. Les enseignants tiennent alors le rôle des élèves et avec l'aide des chercheurs-formateurs analysent les difficultés, erreurs, interventions liées à la tâche. McClain *et al*. (2006) parle de formation de l'enseignant par l'expérimentation de situations présentées aux élèves suivie par une analyse collective des résultats et une discussion sur les réponses, difficultés et conceptions des élèves. Ceci encore une fois avec le soutien des chercheurs-formateurs.

La planification de leçons et le travail en équipe sur des projets où les enseignants tiennent le double rôle d'enseignant et d'apprenant sont d'autres moyens proposés pour la préparation des enseignants. Dans tous les cas, il est important qu'ils puissent expérimenter ce que nous avons appelé le « parcours du statisticien » tout en étant suivis et soutenus par des formateurs qui devraient aussi leur servir de modèle.

Même s'il en a été peu question, il ne faudrait pas négliger dans la formation des enseignants et par conséquent, celle de leurs élèves, l'utilisation d'outils technologiques appropriés. L'ordinateur est désormais un outil indispensable dans le champ de la statistique.

6.2.3 Participation des statisticiens

À ce point, la participation et l'engagement des statisticiens deviennent essentiels. Les statisticiens, comme les mathématiciens avant eux, doivent chercher à prendre une part active dans le développement de l'enseignement de la statistique à l'école. Leur collaboration pour la formation d'enseignants en fonction et de futurs enseignants est nécessaire puisque peu des formateurs actuels ont les compétences adéquates pour donner cette formation (Scheaffer, 2002).

Il serait aussi nécessaire que certains statisticiens s'associent à des enseignants pour les encourager à expérimenter des activités de statistique dans leur classe (McClain *et al*., 2006). Les enseignants ainsi soutenus se sentiront plus confiants et plus à l'aise. Dans l'expérimentation menée par le CIRDIS (Centro Interuniversitario di Ricerca per la Didattica delle Discipline Statistiche, 1998-2000) en Italie, les enseignants participants ont reçu une semaine ou plus de formation (selon le niveau) par les statisticiens et les didacticiens conduisant la recherche. Ils savaient aussi qu'ils pouvaient avoir recours à leur formateur à tout moment pendant l'expérimentation. Ces enseignants ainsi rassurés ont peu fait appel aux chercheurs si ce n'est, à la fin de l'expérimentation, pour relater leur expérience.

Que ce soit pour des activités présentées en classe ou encore dans des manuels, il y a une carence d'exemples de données réelles riches en enseignement pour pouvoir produire des activités significatives (Batanero, 2010). Plusieurs sites[14] en proposent déjà ; cependant, les statisticiens qui rencontrent des cas intéressants dans leur pratique devraient les diffuser et collaborer avec des didacticiens pour en tirer des activités utilisables pour l'enseignement.

Un dernier point est celui de l'évaluation. Tant que les épreuves seront à caractère plutôt mathématique, l'enseignement en sera conditionné. Ceux parmi les statisticiens et les didacticiens de la statistique qui peuvent participer à l'élaboration d'exercices ou d'items d'évaluation doivent faire en sorte que le questionnement demande un raisonnement statistique réel pour éventuellement conduire les enseignants à y préparer leurs élèves.

6.2.4 Collaboration avec la mathématique

Arriver à développer l'enseignement de la statistique à l'école demande une collaboration entre mathématiciens et statisticiens (Moore & Cobb, 2000 ; Estrada & Batanero, 2008). Cette proposition peut soulever une résistance naturelle, mais il faut passer outre afin d'atteindre notre but : le développement de la pensée et du raisonnement statistique chez les élèves par un enseignant spécifique à la statistique.

Cependant, il faut constater que les mathématiciens ont été interpellés par les problèmes d'enseignement depuis longtemps. La didactique de la mathématique s'est développée et est largement reconnue. Elle fait partie de toute formation d'enseignement et il est évidemment impensable d'aller à l'école sans faire de mathématiques.

Une synergie entre les deux disciplines est possible. Ce qui est souhaité pour l'enseignement de la statistique va de pair avec les conceptions plus récentes sur l'enseignement prônées par les didacticiens de la mathématique. L'utilisation de contextes réels et la résolution de problèmes en sont des exemples. La modélisation, souvent au programme de mathématique, pourrait être élaborée avec des données réelles. L'incertitude et la possibilité de résultats différents en analyse de données sont compatibles avec la résolution de problèmes ouverts. En mathématique, on parle de variation (entre deux variables) et en statistique de la variabilité d'une variable : la confrontation entre les deux concepts peut servir à distinguer et approfondir chacun d'eux (Gattuso, 2006, 2008a, 2008b ; Witmer, 2002 ; Burril, 1996).

La statistique peut apporter beaucoup à la mathématique scolaire. Il ne faut pas oublier que l'étude de l'analyse des données permet de travailler de nombreux concepts mathématiques dans un contexte « parlant » et motivant pour l'élève. Non seulement, l'étude de la statistique peut favoriser l'apprentissage de la mathématique, mais elle propose une approche créative qui peut stimuler l'intérêt des élèves pour les « chiffres » et la mathématique. Même si un élève plus fort peut être rebuté par la statistique, un autre moins

[14] Statistique-Canada, Statistix, Census at school, etc.

intéressé par la mathématique peut être intrigué par son expérience en statistique qui lui donne la chance de prendre « un nouveau départ » et lui permet d'envisager de nouvelles perspectives.

7 En guise de conclusion

La tâche est immense et c'est pour cela que toutes les forces en place doivent être sollicitées. Il ne faut pas non plus négliger de regarder ailleurs, de prendre et adapter à nos besoins les expériences des autres et éventuellement de produire de nouveaux outils.

Nous avons peu parlé de l'utilisation de l'informatique. Plusieurs didacticiels, trop souvent en anglais, existent déjà et permettent d'éviter de s'embourber dans des calculs qui entravent l'analyse des données et le développement du raisonnement statistique. Mais même avec un simple chiffrier, les utilisations possibles sont nombreuses. Il va sans dire que, comme nous l'avons mentionné plusieurs fois, l'ordinateur est un support dorénavant indispensable pour le développement du raisonnement statistique même si certaines difficultés y sont associées.

Bien que la formation des maîtres soit trop souvent déficiente pour ce qui est de la statistique, il faut trouver les moyens de pénétrer dans les écoles, d'attirer les enseignants en place et de leur offrir un soutien pour mettre en place la statistique dans le quotidien de la classe.

Références

[1] Aksu, M. (1990), Problem areas related to statistics in training teachers of mathematics in Turkey. In Hawkins, A. (Ed.), *Training Teachers to Teach Statistics. Proceedings of the International Statistical Institute Round Table Conference, Budapest, Hungary, 1988*, International Statistical Institute, Voorburg, Netherlands, 127-137.

[2] Bakker, A. (2001), Symbolizing data into a "bump". In van den Heuvel-Panhizen, M. (Ed.), *Proceedings of the 25th International Conference for the Psychology of Mathematics Education*, Utrecht University, Netherlands, 2, 81-88.

[3] Batanero, C. (2010), Training school teachers to teach statistics: an international challenge, *Statistique et Enseignement*, **1**(1), 5-20, http://www.statistique-et-enseignement.fr/ojs/index.php/StatEns/article/view/3/1 .

[4] Bessant, K. C. and E. MacPherson (2002), Thoughts on the Origins, Concept and Pedagogy of Statistics as a "Separate Discipline", *The American Statistician*, **56**(1), 22-28.

[5] Biehler, R. (2008), From statistical literacy to fundamental ideas in mathematics: How can we bridge the tension in order to support teachers of statistics, *Session plénière, ICMI 2008*, http://www.stat.auckland.ac.nz/~iase/publications/rt08/Panel1_Biehler.pdf .

[6] Burril, G. (1996), Data Driven Mathematics: A Curriculum Strand for High School Mathematics, *The Mathematics teacher*, **89**(6), 460.

[7] Burrill, G. and P. Elliott (2006), Part One: Learning about data and Chance. In Burrill, G. and P. Elliott (Eds.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 1-3.

[8] Canada, D. (2008), United States, Conceptions of Distribution Held by Middle School Students and Preservice Teachers. In Batanero, C., G. Burrill, C. Reading, and A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Conference for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*, http://www.ugr.es/~icmi/iase_study/ .

[9] CIRDIS (1998-2000), *Sperimentazione di nuove strategie didattiche per l'apprendimento della statistica (Experimenting new teaching strategies for learning statistics)*, http://cirdis.stat.unipg.it/index.php?canale=113&lang=eng .

[10] Cobb, G. and D. S. Moore (2000), Mathematics, Statistics and Teaching, *American Mathematical Monthly*, **104**(9), 801-823.

[11] Cobb, P. (1999), Individual and collective mathematical learning: The case of statistical data analysis, *Mathematical Thinking and Learning*, 1, 5-44.

[12] Dunkels, A. (1986), EDA in the primary classroom. Graphing and concept formation combined. In Davidson, R. and J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics*, B. C.: University of Victoria, Victoria, 61-66.

[13] Espinel, M. C., A. Bruno, and I. Plasencia (2008), Statistical Graphs in the Training of Teachers. In Batanero, C., G. Burrill, C. Reading, and A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Conference for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*, http://www.ugr.es/~icmi/iase_study/ .

[14] Estrada, A. and C. Batanero (2008), Explaining teachers' attitudes towards statistics. In Batanero, C., G. Burrill, C. Reading, and A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Conference for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. http://www.ugr.es/~icmi/iase_study/ .

[15] Garfield, J. B. and G. Iddo (1999), Teaching and Assessing statistical reasonning. In Stiff L. V. (Ed.), *Developing Mathematical Reasoning in Grades K-12. 1999 Yearbook of the National Council of Teachers of Mathematics (NCTM)*, VA: NCTM, Reston, 207-219.

[16] Gattuso, L. (2006), Statistics and mathematics: is it possible to create fruitful links?. In Rossman, A. and B. Chance (Eds.), *Proceedings of the 7th International Conference on Teaching Statistics, Salvador, Bahia, Brazil, 2-7 July 2006*, International Statistical Institute, actes sur CD.

[17] Gattuso, L. (2008a), Mathematics in a Statistical Context? In Batanero, C., G. Burrill, C. Reading, and A. Rossman (Eds.), *Joint ICMI/IASE Study Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 Conference and IASE 2008 Round Table Conference, 30 June- 4 July 2008, Monterrey, Mexico,* http://www.stat.auckland.ac.nz/~iase/publications/rt08/T6P1_Gattuso.pdf .

[18] Gattuso, L. (2008b), Statistique et mathématique : y a-t-il vraiment des interactions possibles ? Dans Bednarz, N. et C. Mary (Eds.), *L'enseignement des mathématiques face aux défis de l'école et des communautés. Actes du colloque international Espace Mathématique Francophone Sherbrooke*, Éditions du CRP, actes sur CD.

[19] Gattuso, L. and C. Mary (2001), Pupils perception of the links between data ant their arithmetic average. In van den Heuvel-Panhizen, M. (Ed.), *Proceedings of the 25th International Conference for the Psychology of Mathematics Education*, Utrecht University, Netherlands, 2, 25-32.

[20] Gattuso, L. and C. Mary (2005a), Trois problèmes semblables de moyenne pas si semblables que ça !, *Statistical Education Research Journal*, **4**(2), http://www.stat.auckland.ac.nz/~iase/serj/SERJ4(2)_mary_gattuso.pdf .

[21] Gattuso, L. and C. Mary (2005b), L'influence des grandeurs impliquées sur la résolution d'un problème de moyenne. Dans les *Actes de la rencontre de l'EMF, Espace mathématique francophone, Tozeur, 19-23 décembre 2003*, actes sur CD.

[22] Gattuso, L. and M.-A. Pannone (2000), Une expérimentation d'enseignement des statistiques et les enseignants qui l'ont vécue. Sperimentazione di nuove strategie didattiche per l'apprendimento della statistica cofinanziata dal M.U.R.S.T. e dalle Università di Padova, Palermo, Perugia e Roma « La Sapienza » (1998-2000): *Giornate di Studio, ISTAT, Roma, 6-7 Dicembre 2000*, http://cirdis.stat.unipg.it/files/Sperimentazione/Giornate%20di%20Studio/Relazioni/Gattuso-Pannone.pdf .

[23] Gattuso, L. and M.-A. Pannone (2002), Teacher's training in a statistic teaching experimentation. In *Proceedings of the Sixth International Conference on Teaching Statistics, Durban, South Africa*, International Statistical Institute, 685-692.

[24] Girard, J.-C. (2004), La liaison statistiques-probabilités dans l'enseignement, *Repères-IREM*, 57, 83-91.

[25] Groth, R. (2006), Engaging Students in Authentic Data Analysis. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, National Council of Teachers of Mathematics, Reston, 41-48.

[26] Hawkins, A., F. Jolloffe, and L. Glickman (1992), *Teaching statistical concepts. The effective teachers series*, Longman, London.

[27] Jacobbe, T. (2008), Elementary school teachers' understanding of the mean and median. In Batanero, C., G. Burrill, C. Reading, and A. Rossman (Eds.), *Joint ICMI/IASE Study : Teaching Statistics in School Mathematics. Conference for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*, http://www.ugr.es/~icmi/iase_study/ .

[28] Jelinski, A. (1993), Diagramme circulaire ou orthogonal ? Une efficacité différente des images graphiques dans la transmission de l'information, *Les sciences de l'éducation*, **1**(3), 39-56.

[29] Konold, C. and L. T. Higgins (2003), Reasoning about data. In Kilpatrick, J., W. G. Martin, and D. Schiffer (Eds), *A Research Companion to Principles and Standards for School Mathematics*, VA: NCTM, Reston, 192-215.

[30] London McNab, S. *et al.* (2006), "We Were Nicer, but We Weren't Fairer!" Mathematical Modeling Exploring "Fairness" in Data Management. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 171- 184.

[31] Mary, C. et L. Theis (2007), Les élèves à risque dans des situations problèmes statistiques : stratégies de résolution et obstacles cognitifs, *Revue des sciences de l'éducation*, **33**(2), 579-599.

[32] McClain, K., J. Leckman, and P. Schmitt (2006), Changing the Face of Statistical Data Analysis in the Middle Grades: Learning by Doing. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 229-322.

[33] Meletiou, M. (2003), On the formalist view of mathematics: impact on statistics instruction and learning. In Mariotti, A. (Ed.), *Proceedings of Third European Conference in Mathematics Education, Bellaria, Italy*, European Research in Mathematics Education Society, http://www.dm.unipi.it/~didattica/CERME3/proceedings .

[34] Moore, D. S. and G. Cobb (2000), Statistics and mathematics: Tension and cooperation, *American Mathematical Monthly*, 615-630.

[35] Perelli D'Argenzio, M., S. Rigatti Luchini, G. Moncecchi (2006), Rappresentazioni grafiche nella scuola primaria : uno studio esplorativo / Graphic representation in primary schools : an exploratory analysis, *Induzioni*, 33, F. Serra Editore, http://www.libraweb.net/articoli.php?chiave=2370&rivista=9 .

[36] Queiroz, C. and S. Coutinho (2006), Using experimental approaches for teaching probability: working on a project using face to face and virtual sessions. In Rossman, A. and B. Chance (Eds.), *Proceedings of the 7th International Conference on Teaching Statistics, Salvador, Bahia, Brazil, 2-7 July 2006,* International Statistical Institute, actes sur CD.

[37] Rossman, A., B. Chance, and E. Medina (2006), Some Important Comparisons between Statistics and Mathematics, and Why Teachers Should Care. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 323-335.

[38] Rubel, L. H. (2006), Students' Probabilistic Thinking Revealed: The Case of Coin Tosses. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 49-60.

[39] Russell, S. (1990), Issues in training teachers to teach statistics in the elementary school : a world of uncertainty. In Hawkins, A. (Ed.), *Training Teachers to Teach Statistics. Proceedings of the International Statistical Institute Round Table Conference, Budapest, Hungary, 1988,* International Statistical Institute, Voorburg, Netherlands, 59-71.

[40] Russell, S. (2006), What Does It Mean That "5 Has a Lot"? From the World to Data and Back. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 17-29.

[41] Schield, M. (2001), Statistical Literacy and Mathematic Thinking, http://web.augsburg.edu/~schield/MiloPapers/2000ICME9.pdf.

[42] Scheaffer, R.L. (2002), Statistical Bridges, *Journal of American Statistical Association*, **97**(457), 1-7.

[43] Scheaffer, R. L. (2006), Statistics and mathematics: On making a happy marriage. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 309-321.

[44] Schwartz, L. and D. Whitin (2006), Graphing with Four-Year-Olds: Exploring the Possibilities through Staff Development. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 5-16.

[45] Shaughnessy, M. (2006), Research on Students' Understanding of Some Big Concepts In Statistics. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 77-98.

[46] Shulman, L. S. (1986), Those who understand: Knowledge growth in teaching, *Educational Researcher*, 4-14**.**

[47] Tarr, J., H. Stohl Lee, and R. L. Rider (2006), When Data and Chance Collide: Drawing Inferences from Empirical Data. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of of Mathematics, Reston, 139-150.

[48] Teague, D. J. (2006), Experimental Design: Learning to Manage Variability. In Burrill, G. (Ed.), *Thinking and reasoning with data and chance: NCTM 2006 yearbook*, VA: National Council of Teachers of Mathematics, Reston, 151-170.

[49] Tukey, J. and J. Wilder (1977), *Exploratory Data Analysis*, Addison-Wesley.

[50] Witmer, J. (2002), Data Driven Activities: Teaching Statistics Through Mathematics. In *Proceedings of the Sixth International Conference on Teaching Statistics, Durban, South Africa*, International Statistical Institute, http://www.stat.auckland.ac.nz/~iase/publications/1/2b2_witm.pdf .

Annexe : références pour les enseignants voulant intégrer l'enseignement de la mathématique et de la statistique

Voici les titres de la collection « Data-driven mathematics », White Plains, NY: Dale Seymour :

Burrill, J. (1999), *Mathematics in a world of data.*

Burrill, G., M. Clifford, and R. L. Scheaffer (1998), *Exploring symbols: an introduction to expressions and functions.*

Burrill, G., J. Burrill, P. Hopfensperger, and J. M. Landwehr (1999), *Exploring regression.*

Burrill, G. and P. Hopfensperger (1998), *Exploring linear relations.*

Burrill, G. and P. Hopfensperger (1999), *Exploring systems of equations and in inequalities.*

Burrill, G. and P. Hopfensperger (1999), *Exploring systems of inequalities.*

Burrill, G., J. Witmer, J. Burrill, and J. M. Landwehr (1998), *Advanced modeling and matrices.*

Burrill, J., M. Clifford, and J. M. Landwehr (1999), *Modeling with logarithms.*

Errthum, E. *et al*. (1999), *Exploring projects: planning and conducting surveys and experiments.*

Hopfensperger, P., H. Kranendonk, and R. Scheaffer (1999), *Probability models.*

Hopfensperger, P., H. Kranendonk, and R. Scheaffer (1999), *Probability through data.*

Kranendonk, H. and J. Witmer (1998), *Exploring centers.*