

# STANDARDISATION DES PLANS DE SONDAGE ET ASSURANCE DE QUALITE DANS LES ENQUETES COMPARATIVES

Jean Dumais, Marc Joncas, Sylvie LaRoche<sup>1</sup>

## RÉSUMÉ

Les études comparatives internationales en éducation ont éprouvé le besoin de normaliser leurs méthodes statistiques, en particulier, les plans de sondage. Nous examinons les bénéfices et défis liés à la mise en place d'un plan de sondage normalisé pour ces études. Nous terminons en discutant des leçons apprises au cours de plus de quinze ans de participation à différentes enquêtes internationales.

MOTS CLÉS : Études internationales; Normes; Échantillonnage.

## ABSTRACT

International comparative studies in education have standardised their statistical methods, in particular, the sampling design they use. Benefits and challenges associated with such standardised designs will be examined. Lessons learned during more than fifteen years of practice will be given in closing remarks.

KEY WORDS: International studies; Standards; Sampling.

## 1. INTRODUCTION

Les études comparatives à caractère international dans le domaine de l'éducation existent depuis de nombreuses années. Ces enquêtes visent à mesurer l'efficacité des systèmes éducatifs dans leur ensemble et s'adressent généralement aux élèves et/ou aux enseignants. Plusieurs de ces enquêtes évaluent les compétences acquises par les élèves et permettent ainsi de comparer les performances des systèmes d'éducation en place dans les pays participants. Parmi les plus connues, mentionnons le Programme international pour le suivi des acquis des élèves (PISA<sup>2</sup>), le Programme international de recherche en lecture scolaire (PIRLS<sup>3</sup>) et les Tendances de l'enquête internationale sur la mathématique et les sciences (TIMSS<sup>4</sup>). Depuis 2008, l'Organisation de coopération et de développement économiques (OCDE) a aussi mise sur pied l'enquête internationale sur les enseignants, l'enseignement et l'apprentissage (TALIS) qui s'intéresse aux convictions, pratiques et attitudes des enseignants, à leurs conditions de travail et à leur environnement pédagogique.

Comme la plupart de ses études établissent un classement des systèmes d'éducation des pays participants basé sur la performance des élèves, leurs résultats sont hautement médiatisés et, parfois, politiquement délicats. De par leur nature comparative, les études internationales demeurent donc sujettes à la controverse. Ainsi, les personnes responsables de ces études font face à de nombreux défis afin de s'assurer que tous les aspects des enquêtes soient contrôlés et vérifiés de façon à ce que les résultats en découlant soient comparables et crédibles.

---

<sup>1</sup> Jean Dumais, Statistique Canada, Ottawa (ON), Canada, K1A 0T6 ([jean.dumais@statcan.gc.ca](mailto:jean.dumais@statcan.gc.ca)); Marc Joncas, Statistique Canada, Ottawa (ON), Canada, K1A 0T6 ([marc.joncas@statcan.gc.ca](mailto:marc.joncas@statcan.gc.ca)); Sylvie LaRoche, Statistique Canada, Ottawa (ON), Canada, K1A 0T6 ([sylvie.laroche@statcan.gc.ca](mailto:sylvie.laroche@statcan.gc.ca)).

<sup>2</sup> *Programme for International Student Assessment* menée par l'Organisation de coopération et de développement économiques (OCDE).

<sup>3</sup> *Comparison of the Progress in International Literacy Study*. Cette étude est réalisée par le TIMSS&PIRLS International Study Center et financée par Association internationale pour l'évaluation du rendement scolaire (IEA - *International Association for the Evaluation of Educational Achievement*).

<sup>4</sup> *Trends in International Mathematics and Science Study*. Cette étude est également réalisée par ISC et financée par IEA.

**Exhibit 2.6 Trends in Average Achievement in Advanced Mathematics by Gender**

Country	Females		Males	
	2008 Average Scale Score	1995 to 2008 Difference	2008 Average Scale Score	1995 to 2008 Difference
Italy	454 (9.3)	-23 (15.7)	446 (8.3)	-41 (15.2) ▼
Russian Federation	551 (7.7)	25 (11.4) ▲	569 (7.4)	0 (11.3)
‡ Slovenia	448 (5.3)	-21 (12.5)	472 (4.3)	-14 (11.9)
Sweden	404 (6.9)	-88 (8.5) ▼	418 (6.1)	-88 (9.6) ▼

▲ 2008 average significantly higher than 1995  
▼ 2008 average significantly lower than 1995

‡ In 1995, did not satisfy guidelines for sample participation rates (see Appendix A).  
( ) Standard errors appear in parentheses.

SOURCE: IEA TIMSS Advanced 2008 ©

**Figure 1: Exemple de tableau publié par TIMSS Advanced 2008**

L'échantillonnage est certainement une composante importante à cet égard. La méthodologie d'échantillonnage d'enquêtes antérieures a souvent été l'objet de critiques. Nous pouvons certainement nous poser les questions suivantes : est-ce l'échantillon d'un pays est représentatif de la population cible? Est-ce que les niveaux d'exclusions sont comparables d'un pays à l'autre? Les échantillons ont-ils été sélectionnés de façon appropriée? Est-ce que les taux de participation sont acceptables? Ces questions et bien d'autres ont donné naissance à une série de contrôles et de standards propres aux études comparatives dans le domaine de l'éducation, lesquels font l'objet d'une revue dans les sections subséquentes. Cette revue est limitée aux aspects touchant particulièrement le plan de sondage. Dans un premier temps, nous décrivons les conditions définissant un contexte idéal menant à un plan de sondage adéquat, pour naturellement aborder les difficultés inhérentes à l'implantation d'un tel plan sur le terrain. Cette section est suivie d'une courte discussion sur les raisons qui nous amènent à standardiser dans un contexte international. On donne ensuite une série d'exemples de contrôles et de standards parmi les plus couramment utilisés dans ce genre d'enquête avant de conclure en mentionnant les leçons retenues après presque vingt ans d'expérience à travailler sur les enquêtes comparatives internationales en éducation.

## 2. CONTEXTES

### 2.1 Contexte idéal

Dans un contexte idéal, un plan de sondage devrait mener à des résultats sans biais, exacts et comparables internationalement. Pour ce faire, trois conditions apparaissent incontournables lors de la conception d'enquêtes internationales.

D'abord, on veut la meilleure couverture possible de la population cible ; idéalement, la population d'enquête doit correspondre à la population cible. Cette condition est primordiale étant donné que les inférences se rapportent à la population d'enquête. Une fois la collecte terminée, il n'est pas toujours possible de faire des ajustements lors de l'estimation afin de corriger les défauts de couverture. Il est souvent plus facile de faire comprendre l'importance de couvrir entièrement la population d'intérêt lorsqu'il s'agit d'un recensement. Que vaudrait un recensement avec une couverture de seulement 80 % ?

Dans le cas d'une enquête-échantillon, plusieurs pensent, à tort, que cette condition n'est pas aussi capitale. Dans un contexte international, ceci est encore plus important puisque que c'est le premier aspect pouvant miner la crédibilité de l'enquête. Des taux de couverture différents d'un pays à l'autre ouvrent inévitablement la porte à des discussions remettant en cause la validité des comparaisons et peuvent ainsi limiter les possibilités d'analyse. Dans le contexte des enquêtes touchant l'éducation, cet élément est particulièrement important puisqu'il peut laisser l'impression que les résultats seront biaisés si la population non couverte d'un pays incluent les étudiants les moins performants.

Ensuite, un plan d'échantillonnage valide est la deuxième condition essentielle. Il constitue notre lien avec la population d'enquête. Toutes les unités de la population d'enquête se doivent d'avoir une chance (probabilité d'inclusion non nulle) d'être choisies sous peine de voir le niveau d'exclusions s'accroître (et ainsi retomber avec une couverture non complète). Cette probabilité d'inclusion doit être connue et calculable afin d'exclure un risque de biais. Il est important aussi d'assurer une mise en œuvre adéquate de ce plan d'échantillonnage (car il ne sert à rien de faire de beaux plans d'échantillonnage si ceux-ci ne sont pas exécutés correctement sur le terrain). À défaut de disposer d'un plan d'échantillonnage valide, il devient hasardeux d'inférer les résultats de l'enquête à la population d'enquête.

Finalement, comme troisième condition, les erreurs d'échantillonnage devraient être les plus petites possibles. Les résultats d'enquête seraient donc le plus près possibles des valeurs qui auraient été obtenues d'un recensement.

Dans un contexte idéal où toutes ces conditions seraient réunies, avec une mise en œuvre parfaite du plan d'échantillonnage, des procédures de collecte, et des taux de réponse de 100%, le besoin d'établir des standards tels que définis à la section 4.1 n'existerait pas.

## **2.2 Contexte international (sur le terrain)**

Dans un projet d'enquête regroupant plusieurs pays, il est rare de retrouver un contexte idéal dans lequel les conditions discutées à la section précédente sont pleinement respectées. Les particularités géographiques ou culturelles, la disponibilité et l'accès aux données administratives pertinentes, la capacité de mise en œuvre d'enquêtes, le fardeau de réponse, la culture d'enquêtes<sup>5</sup> et les systèmes d'éducation eux-mêmes sont autant de facteurs qui différencient les pays les uns des autres. Invariablement, les taux de couverture, de réponse, la qualité de la mise en œuvre, bref tout ce qui touche au plan de sondage sera affecté à des degrés différents d'un pays à l'autre. Ce constat n'implique pas que toute tentative de comparaison entre les pays soit vouée à l'échec, mais que pour ce faire, il est nécessaire d'établir des normes et des standards minimaux en deçà desquels les comparaisons deviennent douteuses.

## **3. POURQUOI STANDARDISER**

La mise en œuvre de standards et de contrôles dans ce genre d'études permet de valider la comparaison des résultats et renforce la crédibilité de ceux-ci aux yeux des utilisateurs. Le but recherché est de pouvoir attribuer une différence (statistiquement) significative observée dans les résultats à une réelle différence dans les populations comparées, et non au fruit d'une combinaison d'erreurs non contrôlées (taux de réponse inadéquat, couverture mal assurée, piètre qualité des mises en œuvre des procédures sur le terrain, erreurs de mesures, *etc.*). La motivation des utilisateurs à procéder à des analyses des résultats est fortement influencée par la qualité des données et la pertinence des comparaisons de résultats.

L'établissement de standards facilite également la mise en œuvre du plan de sondage. Dans le domaine particulier des enquêtes internationales touchant l'éducation, la standardisation se traduit en l'unification des procédures. À titre d'exemple, on notera que le plan d'échantillonnage, la taille d'échantillon, le mode de collecte, ou encore le libellé des questionnaires d'enquêtes prestigieuses telles que PISA ou TIMSS est à peu de choses près le même pour tous les participants (des exemples de standards sont donnés dans les prochaines sections). Cette unification permet un meilleur équilibre des charges de travail de chacun des pays participants à l'enquête et permet aussi la mise sur pied de mesures minimales efficaces et uniformisées de contrôle de la qualité des données.

En résumé, l'établissement de standards et de contrôles est profitable à tous les intervenants. Aux commanditaires qui seront rassurés du bien-fondé de leur investissement en obtenant des résultats valables et comparables, aux gestionnaires d'enquêtes qui peuvent se porter garants de la qualité des procédures et de la validité des résultats, et enfin aux pays participants qui obtiennent des résultats d'une qualité assurée et ce pour une charge de travail somme toute équivalente peu importe les conditions, contraintes et environnement qui leurs sont propres.

## **4. EXEMPLES DE STANDARDS ET CONTROLES**

### **4.1 Introduction**

La différence que nous faisons entre standards et contrôles est la suivante : les standards sont des normes établies pour assurer la qualité des données. Ces normes sont décrites et documentées. La non-conformité aux standards se traduit habituellement par des

---

<sup>5</sup>On entend par culture d'enquête l'ouverture qui existe face aux enquêtes dans les différents pays. Pour certains, la participation aux enquêtes peut être obligatoire alors que pour d'autres les enquêtes sont volontaires et parfois perçues comme étant une atteinte à la vie privée.

conséquences pour les pays fautifs, allant de simples remarques dans les tableaux de publication à la mise en annexe<sup>6</sup> de tous leurs résultats. Le caractère des contrôles est quant à lui moins formel et se manifeste par l'application de critères plus souples. On parle plutôt de procédures de contrôle de qualité. Ces procédures de contrôle ont été définies et mises en application avec l'expérience acquise au cours des années. Cependant, certains contrôles qui existaient au début des enquêtes comparatives internationales sont maintenant devenus des standards établis.

## 4.2 Populations cible et d'enquête

Les premiers exemples de standards se rapportent à la population cible et la population d'enquête. Ainsi, chaque pays participant est tenu de fournir une description de sa population nationale. Il doit au minimum, décrire le système éducatif (par exemple, l'âge minimal d'entrée à l'école, une description de la structure scolaire, le niveau ISCED<sup>7</sup>). La population d'enquête doit couvrir au moins 95 % de la population cible (ce standard est probablement le plus reconnu de la communauté en éducation). Toute divergence entre la population d'enquête et la population cible doit être documentée (nature et ampleur des exclusions). Finalement, une couverture de la population cible pour un pays donné de moins de 95 % entraînera automatiquement une annotation dans les publications internationales. Cette norme de 95 % peut sembler élevée mais il est important de se rappeler que dans les publications internationales, les tableaux de résultats comportent une ligne par pays. L'utilisateur suppose naturellement que les résultats sont représentatifs du pays dans son ensemble et c'est donc sur cette base même que les comparaisons entre les pays sont faites. Tel que mentionné précédemment à propos des enquêtes en éducation, il y a souvent une forte présomption de corrélation entre la couverture et la performance mesurée. On peut difficilement réduire ces exigences et encore prétendre que les comparaisons sont crédibles.

Certains contrôles sont fréquemment utilisés pour assurer que la définition de la population cible nationale corresponde à la définition de la population cible internationale. Des vérifications additionnelles de l'information fournie par le pays (tel que l'âge, le nombre d'années de scolarisation, la fréquentation scolaire<sup>8</sup>, *etc.*) sont effectuées en consultant des sources externes de données. Dans le contexte d'enquêtes répétées, nous vérifions que la définition des populations est comparable d'un cycle à l'autre pour permettre des estimations de tendances valides. Si la population cible change d'un cycle à l'autre, il faut identifier la portion de la population qui est commune aux deux cycles pour pouvoir effectuer l'analyse des tendances. Cette situation peut survenir suite à un changement dans le système d'éducation d'un pays qui ferait en sorte que l'année scolaire représentant 4 années d'éducation formelle (servant souvent de base pour définir une population) passe de la 4<sup>ème</sup> année à la 5<sup>ème</sup> année avec la venue du nouveau cycle (s'échelonnant sur 3 à 5 ans en général). Une réforme peut également modifier l'âge d'entrée à l'école. Présentement, il n'y a pas à notre connaissance de normes stipulant de manière explicite un pourcentage à atteindre dans ces situations (tel qu'un pourcentage minimum de la population commune aux deux cycles).

## 4.3 Plan de sondage

Nous discuterons ici des standards appliqués à la base de sondage, au mode de sélection des échantillons, à la taille des échantillons et à la mise en œuvre du plan d'échantillonnage.

### 4.3.1 Base de sondage

Dans les enquêtes en éducation, la base de sondage utilisée est souvent composée d'une liste d'écoles et la mesure de taille correspond aux nombres d'unités faisant partie de la population cible (professeurs ou élèves). À notre connaissance, il n'existe pas de standards établis et publiés pour en vérifier la qualité. Il y a par contre de bonnes pratiques (contrôles) mises en place qui correspondent aux vérifications habituellement faites sur toutes bases de sondage. Dans un premier temps, nous vérifions que la base de sondage soit la plus à jour possible. Ensuite, on s'assure que la base de sondage fournie par les participants donne une couverture complète de la population d'enquête, et qu'elle ne contienne pas de données erronées, de duplicatas, ou d'éléments extérieurs à la population d'enquête/cible. De plus, autant que possible, une mesure de taille à jour pour chacune des unités constituant la base est exigée. Nous insistons également auprès des pays pour que la base de sondage fournie donne accès à l'ensemble de la population cible. Ceci permet de mieux estimer et documenter les exclusions. Enfin, nous utilisons des outils tels que le web, les informations des cycles précédents, les informations provenant d'autres pays pour valider les informations fournies par les représentants des pays. Par exemple, plusieurs pays ont des écoles internationales. Un pays en particulier peut avoir omis d'inclure ces écoles parce qu'il ne les considère pas comme des écoles faisant partie de son système d'éducation.

### 4.3.2 Mode de sélection

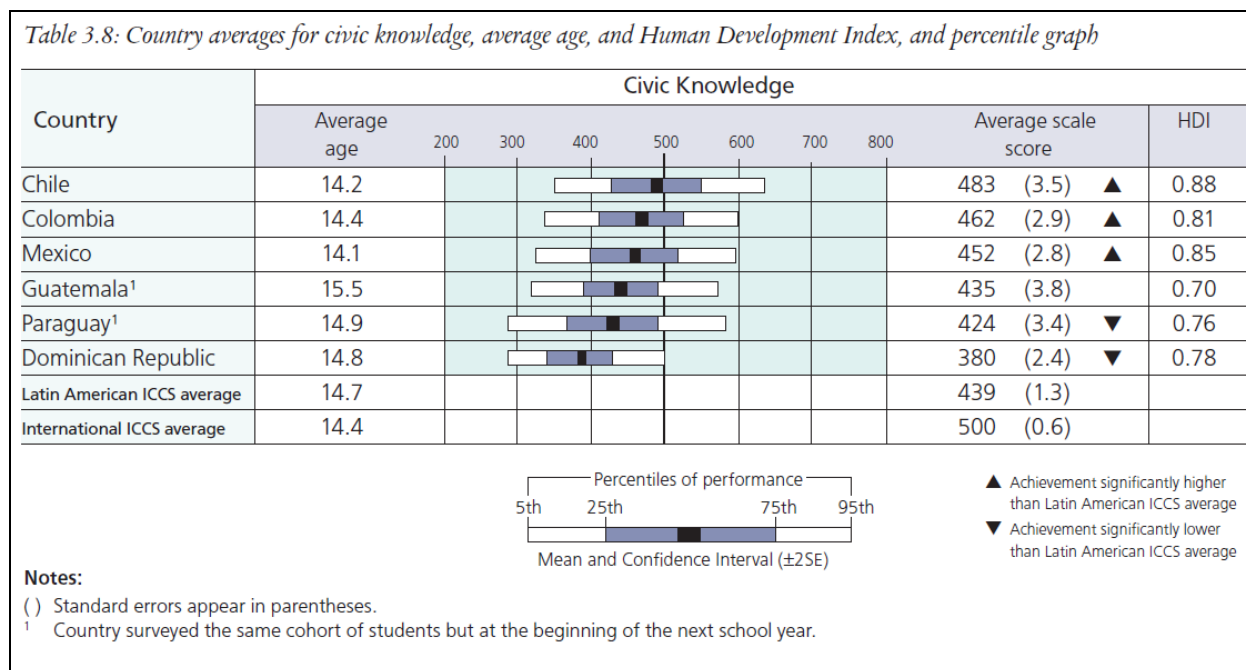
---

<sup>6</sup> Voir l'exemple donné en introduction. Les résultats des pays fautifs sont exclus des tableaux principaux et se retrouvent en annexe à la fin des rapports.

<sup>7</sup> L'*International Standard Classification of Education* (ISCED) est une norme de classification des systèmes d'éducation provenant de l'UNESCO.

<sup>8</sup> La fréquentation scolaire se définit comme le pourcentage de la cohorte d'âge d'une année scolaire donnée qui fréquente l'école.

S'agissant du mode de sélection des échantillons, la norme est d'imposer un mode de sélection unique pour tous les pays participants. Les adaptations ou déviations sont permises mais elles doivent être approuvées par les gestionnaires de l'enquête avant la mise en œuvre et elles doivent aussi être documentées. L'imposition d'un mode de sélection unique nous permet de mettre au point et d'utiliser des programmes de sélection et de pondération des échantillons utilisables pour tous les pays, minimisant ainsi les risques d'erreurs. Cette approche facilite la répartition équitable des charges de travail pour chacun des participants, permettant d'uniformiser considérablement les tailles d'échantillons d'un pays à l'autre. Elle facilite également la validation des sélections et permet de minimiser le nombre de programmes de contrôle nécessaires pour la mise en œuvre. Enfin, l'adoption d'un mode unique de tirage permet d'uniformiser les opérations de collecte et donc de limiter le nombre de manuels d'opérations requis. Ainsi, les risques d'erreurs sont réduits, évitant que des différences dans les instructions aient des répercussions sur la qualité et la comparabilité des données. Enfin, le recours à un seul mode de sélection contribue à rassurer les participants sur la comparabilité des résultats : on s'attend à ce que les erreurs non dues à l'échantillonnage soient comparables. On s'attend également à ce que les erreurs d'échantillonnage soient du même ordre de grandeur. Notons qu'il y a des conséquences à ne pas satisfaire aux standards. Le risque que les données ne soient pas publiées ou qu'elles soient annotées dans les tableaux croît considérablement si le plan n'est pas approuvé ou que des anomalies sont observées. (Voir la note au bas du tableau présenté à la Figure 2.)



**Figure 2: Exemple de tableau publié par l'enquête *International Civic and Citizenship Education Study 2009***

### 4.3.3 Taille d'échantillon

Invariablement dans ce genre d'enquête, la taille minimale de l'échantillon est normalisée. La plupart du temps, celle-ci est fixée en fonction des marges d'erreur souhaitées et des contraintes intrinsèques à l'étude. Un autre standard souvent retenu concerne l'identification a priori d'écoles dites de remplacement. De façon générale, il y a un maximum de deux écoles de remplacement pour chaque école originalement choisie. Encore une fois, toute déviation doit être documentée et approuvée. Notons que les écoles de remplacement ne peuvent être utilisées que pour remplacer les écoles éligibles mais refusant de participer. L'utilisation de remplacement permet de satisfaire aux exigences concernant la taille des échantillons et peut contribuer à minimiser le risque de biais. Nous conservons cependant des exigences strictes en ce qui a trait au taux minimal de participation des écoles originalement sélectionnées (voir la section suivante).

### 4.3.4 Mise en œuvre

Le standard le plus reconnu en ce domaine est l'obligation pour tous les pays de participer à une mise à l'essai. Celle-ci permet de tester les procédures sur le terrain et surtout de mettre en place des correctifs (comme c'est souvent le cas) avant la mise en œuvre de l'enquête. Cette participation à la mise à l'essai est en principe obligatoire sous peine de voir les résultats des pays fautifs exclus des publications internationales.

Un autre standard reconnu est le taux de réponse minimal. Nous pouvons en général distinguer trois zones :

- 1) La zone du minimum absolu ou zone rouge. À défaut de ne pas satisfaire à ces taux minimum, un pays se verra exclu de toutes les publications internationales ;
- 2) À l’opposé, il y a la zone verte. Un pays se situe dans la zone verte si ses taux de participation sont au-delà d’un certain seuil, et ce sans le recours aux écoles de remplacement (p.ex. 85% pour TIMSS). On dit alors que le risque de biais dans les résultats découlant des données provenant de ce pays est négligeable. Si ce seuil est atteint seulement une fois que les écoles de remplacement sont prises en ligne de compte, alors les résultats de ce pays sont inclus dans les publications internationales mais comprennent une annotation afin d’alerter les utilisateurs d’un risque accru de biais ;
- 3) Finalement, il y a la zone grise. Si les taux de participation d’un pays sont au-delà du seuil minimum défini par la zone rouge, mais n’atteignent pas le seuil minimum défini par la zone verte, même après le recours aux écoles de remplacement. Les résultats de ce pays peuvent se retrouver en fin de tableaux, en annexe, voire ne pas être publiés ; la décision est prise au cas par cas.

Il est impératif que les taux de réponse (parfois aussi appelés taux de participation) soient documentés afin que les analystes puissent évaluer la qualité des inférences et analyses issues des données.

En ce qui touche les contrôles, notons que durant la mise en œuvre, les pays ont souvent comme instruction de contacter les gestionnaires des enquêtes lorsqu’ils rencontrent une situation inhabituelle. Ceci permet de contrôler et d’agir avant que la collecte de données ne soit terminée et qu’il n’y ait plus de correctifs possibles. La présence d’une mesure de la taille des écoles sur la base de sondage permet un autre contrôle qui consiste à comparer cette dernière à la taille observée une fois sur le terrain. Il est alors possible de demander des justifications et une documentation plus pointue pour les cas où les différences sont importantes (omission de classes, erreur dans l’identification de l’école, changement dans la structure même de l’école visée, *etc.*). De plus, il n’est pas rare de procéder à une validation du statut des écoles non participantes une fois la collecte terminée afin de déterminer s’il n’a pas lieu de considérer un statut plus approprié (certains refus comme des exclusions par exemple). Toujours dans le domaine des contrôles, les erreurs-types sont calculées entre autres pour détecter les valeurs aberrantes, les valeurs influentes, les poids influents ou aberrants en fonction de la principale variable d’intérêt. Il est également possible d’effectuer une comparaison des estimations observées et attendues (comme par exemple, les taux d’exclusions à l’intérieur des écoles comparés à ceux des cycles précédents, les totaux de populations comparés aux totaux connus des cycles précédents, *etc.*). Tous ces contrôles nous permettent de détecter de possibles manquements aux règles dictées par le plan d’échantillonnage. Des justifications écrites auprès des participants sont habituellement exigées pour toutes anomalies détectées.

Comme dernier point ici, nous aimerions souligner l’importance d’effectuer une évaluation de la mise en œuvre. La revue des plans de sondage et de la mise en œuvre est en général effectuée en présence d’un expert extérieur au cercle de gestion de l’enquête. Cette évaluation et approbation indépendante apporte une crédibilité importante à l’enquête. Finalement, il est essentiel de clore le tout par la rédaction d’un rapport technique décrivant l’ensemble des procédures touchant le plan de sondage et sa mise en œuvre.

## 5. CONCLUSION

De notre expérience sur les enquêtes comparatives internationales, nous retenons ceci :

- 1) qu’il est difficile d’avoir des standards qui à la fois répondent à tous les besoins tout en conservant une certaine souplesse. À chaque enquête ou cycle, nous devons faire face à des situations particulières. Il devient donc important de mettre en place une équipe technique responsable de fournir un soutien aux participants et d’inviter ceux-ci à consulter cette équipe avant et pendant la mise en œuvre de l’enquête afin de palier les différents imprévus ;
- 2) que l’établissement de standards est nécessaire et primordial afin de dissiper tous doutes sur la pertinence des analyses découlant de l’enquête ;
- 3) qu’il est important d’exercer à un coût raisonnable un contrôle sur toutes les procédures menant à bonifier la qualité des données recueillies ;
- 4) qu’il est tout aussi important de quantifier et de documenter les actions prises afin d’assurer la qualité des données et leur comparabilité et ainsi renforcer la confiance envers les chargés de la mise en œuvre.

On peut certes prétendre qu’il y a toujours place à amélioration. Certains contrôles décrits plus haut pourraient facilement être raffermis pour devenir des standards. Par contre, comme gestionnaires d’enquêtes, on se doit de se montrer prudent et de conserver

une certaine souplesse, toujours avec l'optique de garantir un niveau de qualité acceptable. Les standards sont contraignants pour les pays participants. Donc, devrait-on viser plus de standards au détriment d'une certaine souplesse et d'accommodements sur le terrain, ou plutôt son contraire, avec les risques de dérives possibles? La discussion reste ouverte...

## RÉFÉRENCES

International Association for the Evaluation of Educational Achievement (2011) *ICCS 2009 Latin American Report : Civic knowledge and attitudes among lower-secondary students in six Latin American countries*, Amsterdam: IEA.

OECD (2012) *PISA 2009 Technical Report*, PISA, OECD Publishing, <http://dx.doi.org/10.1787/9789264167872-en>

TIMSS & PIRLS International Study Center (2009), *TIMSS Advanced 2008 International Report*, Chestnut Hill, MA; TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center (2007), *TIMSS 2007 Technical Report*, Chestnut Hill, MA; TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center (2005), *TIMSS 2007 school sampling manual*, Chestnut Hill, MA; TIMSS & PIRLS International Study Center, Boston College.