

Evaluation externe d'un lycée-pilote au Luxembourg – utilisation de questionnaires en miroir pour comparer l'opinion des différents acteurs de la communauté scolaire

Denise Villányi, denise.villanyi@gmx.de, Université du Luxembourg

Giovanna Mancuso, giovanna.mancuso@uni.lu, Université du Luxembourg

Débora Poncelet, debora.poncelet@uni.lu, Université du Luxembourg

Romain Martin, romain.martin@uni.lu, Université du Luxembourg

Résumé : Dans cette communication les résultats issus de l'évaluation externe d'un lycée-pilote au Luxembourg seront présentés. L'objectif de cette évaluation était d'établir un bilan après cinq ans de fonctionnement de ce lycée et de donner un feedback formatif à la communauté scolaire sur les pratiques d'apprentissage, d'évaluation et d'autoévaluation. Des questionnaires en miroir ont été construits pour permettre la mise en évidence et la comparaison des perspectives des différents acteurs de cette communauté scolaire (équipe éducative: équipe pluri-professionnelle, élèves et parents). Pour mener à bien cette évaluation, en nous donnant la possibilité de comparer toutes les perspectives des différents groupes évalués, nous avons cherché un outil pouvant nous fournir des réponses dans un continuum métrique plutôt qu'ordinal. Nous avons utilisé une échelle visuelle analogique (EVA) pour mesurer les perceptions subjectives des répondants et pour nous permettre de faire les comparaisons voulues. L'objet de cette communication est double, le croisement des différentes opinions des groupes de la communauté scolaire en ce qui concerne les pratiques d'apprentissage, d'évaluation et d'autoévaluation de ce lycée mais aussi la présentation d'un outil d'évaluation métrique peu connu dans les sciences de l'éducation.

Mots-clés : évaluation scolaire, évaluation/autoévaluation, questionnaires en miroir, échelle visuelle analogique (EVA)

1. Introduction

Il existe peu de littérature sur les recherches en science sociale s'intéressant aux concordances entre différents types d'informateurs (père, mère, enfant, enseignant, observateurs extérieurs...). Les études existantes en psychologie clinique et en santé médicale s'intéressant aux concordances entre différents types d'informateurs pour mesurer des construits psychologiques et des comportements montrent qu'elles sont souvent faibles. (Dierendonck, 2010) Dans le cadre de l'évaluation d'un établissement scolaire, les chercheurs-évaluateurs doivent s'attendre à des points de vue hétérogènes des différents groupes (direction, équipe éducative, élèves, parents) sur l'organisation et le fonctionnement de leur école. En accord avec la conclusion de Dierendonck (2010) précisant que « les chercheurs, qu'ils travaillent dans le domaine de la santé mentale ou en éducation, ont tout intérêt à collecter les différents points de vue des intervenants pour approcher au mieux la réalité qu'ils étudient » (p.84), nous voudrions souligner l'importance d'impliquer tous les acteurs concernés lors de l'évaluation d'un projet ou d'un établissement scolaire.

Les résultats de la recherche présentés dans cette communication sont issus d'un projet d'évaluation du lycée-pilote «Lycée Ermesinde» (LE), commandité par le Ministère de l'Education nationale et de la Formation professionnelle (MENFP) au Luxembourg. L'évaluation du LE avait été commanditée avec les objectifs :

- (a) d'établir un bilan après cinq ans de fonctionnement de l'école ;
- (b) de donner un feedback formatif à la communauté scolaire.

Pour mieux comprendre les observations que nous avons recueillies, nous allons dans un premier temps vous décrire quelques caractéristiques de ce lycée-pilote.

Le LE a ouvert ses portes en septembre 2005 à Luxembourg, pour accueillir progressivement des élèves du cycle inférieur des différents régimes d'études de l'enseignement secondaire. Entre 2005 et 2011, l'école s'est élargie chaque année d'une promotion. Actuellement, il est possible de faire un diplôme de fin d'études secondaires (régime classique) au Lycée Ermesinde. L'école a été reconnue comme lycée régulier au cours de l'année 2011. Son nom a été changé de « Neie Lycée » (nouveau lycée) en Lycée Ermesinde.

Le LE se caractérise par:

- un régime « à plein temps » de type lié (obligatoire vs. optionnel) dont l'objectif est la favorisation d'un encadrement individualisé des enfants facilitant la réussite scolaire pour tous ;
- la mise en œuvre d'un enseignement et d'un encadrement éducatif intégrés des élèves, ceci par une équipe pluri-professionnelle composée d'enseignants, d'éducateurs gradués, de spécialistes (artisans et artistes offrant des activités spécialisées) ainsi que d'un psychologue et d'un assistant social ;
- un tutorat hebdomadaire et individuel. Chaque élève est accompagné dans son parcours et orientation scolaire par un tuteur individuel, qu'il voit chaque semaine pour un entretien de 15 minutes environ. Ce tuteur est également la personne de contact privilégiée entre école et famille ;
- la recherche d'une coopération étroite entre école et famille (information, communication et participation) en vue d'un partenariat éducatif entre école et famille ;
- une approche par compétences en ce qui concerne l'apprentissage, l'enseignement et l'évaluation. En conséquence, l'évaluation formative et l'autoévaluation de l'élève ont une place importante dans le setting d'apprentissage.

L'objet de cette communication est avant tout de souligner l'importance de questionner tous les acteurs d'un système lors d'une évaluation pour éviter les croyances provoquant des distorsions entre groupes. Une solution pour éviter cette distorsion est l'utilisation d'un questionnaire en miroir.

Nous allons illustrer les avantages de l'utilisation de ces questionnaires en miroir en vous montrant une partie restreinte des résultats obtenus lorsque nous croisons différentes opinions des groupes de la communauté scolaire sur les pratiques d'apprentissage, d'évaluation et d'autoévaluation.

Une difficulté dans la création des questionnaires était d'adapter le vocabulaire à tous les groupes cibles. Par exemple : Les membres de l'équipe éducative sont habitués au vocabulaire technique de leur profession alors que les parents ne le sont pas. Une autre difficulté était de trouver un instrument de mesure permettant de faire la comparaison entre les différents groupes et de donner la possibilité statistique de croiser plusieurs variables.

Le choix s'est porté sur l'échelle visuelle analogique (EVA), un instrument de mesure qui recueille des données quantitatives et permet de faire des analyses de variance (ANOVA). L'EVA est un outil de mesure qui capture des données subjectives et permet de quantifier ces données sur un continuum métrique.

Cette échelle est surtout utilisée dans les domaines des sciences médicales et psychologiques: des milliers d'articles publiés ont utilisé cet outil alors que seuls quelques dizaines d'articles publiés dans le domaine des sciences de l'éducation l'ont utilisé. Nous la retrouvons surtout dans les recherches abordant la formation des médecins, des infirmiers et des agents soignants (A'Campo, Spliethoff-Kamminga & Macht 2010 ; Turner, van de Leemput, Draaisma, Oosterveld & ten Cate 2008 ; Foley 2006), son utilisation est rare dans le domaine de l'éducation scolaire (Myford 2002 ; Bebell, O'Dwyer, Russel & Hoffmann 2010). Récemment, Poncelet et ses collègues (*accepté*) ont utilisé une méthode mixte de recherche afin d'évaluer l'impact d'une recherche-action de deux années au sein d'une école primaire luxembourgeoise. Exploitant des données quantitatives d'un questionnaire par le

biais des EVA, ils ont enrichi cette procédure en la couplant avec un entretien directif basé sur ce même outil.

L'EVA a été comparée avec d'autres types d'échelles pour déterminer si elle produit des résultats similaires.

- Avec les échelles de Borg et de Likert pour mesurer l'essoufflement et la fatigue des participants (Grant, Aitchison, Henderson, Christie, Zare, Mc Murray & Dargie 1999). L'EVA était supérieure à la fois pour la reproductibilité et la sensibilité des réponses.
- Avec des échelles de Likert pour évaluer la capacité d'imagerie mentale (Quilter, Band & Miller 1999). Ces auteurs ont constaté que les deux instruments fournissent des résultats similaires.

L'objet de cette communication est donc double : présenter les avantages d'un questionnaire en miroir pour connaître les différentes opinions des groupes de la communauté scolaire en ce qui concerne les pratiques d'apprentissage, d'évaluation et d'autoévaluation de ce lycée, mais aussi vous présenter un outil d'évaluation métrique peu connu dans les sciences de l'éducation.

2. Méthodologie

La recherche a été menée en deux phases :

- une phase exploratoire moyennant une analyse des documents, des observations sur le terrain et des entretiens individuels afin de connaître le contexte de l'école ;
- une deuxième phase ayant comme but l'évaluation du fonctionnement du lycée en impliquant les perspectives de la communauté scolaire : celles de l'équipe éducative (équipe pluri-professionnelle), des élèves et des parents.

Lors de la deuxième phase, trois questionnaires en miroir ont été construits pour recueillir les points de vue des différents groupes faisant partie de la communauté scolaire (enseignants, éducateurs gradués, professionnels assurant les activités complémentaires, élèves et parents). Les observations et entretiens de la première phase de recherche ont permis de construire des items « sur mesure », reprenant les pratiques et situations concrètes pour les membres de la communauté scolaire. Les questions ont été adaptées selon le groupe ciblé. Chaque questionnaire comprenait:

- une partie descriptive avec des données sociodémographiques ;
- une partie avec des items sur le fonctionnement et la satisfaction concernant différents domaines du quotidien scolaire dont les pratiques d'apprentissage, l'évaluation et l'autoévaluation.

Pour recueillir les différents points de vue, nous avons adopté pour la plupart des items du questionnaire des EVA. Ci-dessous, nous vous présentons un exemple d'item sous forme d'échelle visuelle analogique:

Etes-vous satisfait du tutorat tel que pratiqué au Neie Lycée ?

**Pas du tout
d'accord**



**Tout à fait
d'accord**

Par exemple, pour répondre à la question « Etes-vous satisfait du tutorat tel que pratiqué au Neie Lycée? », nous présentons aux participants une ligne continue de 10 cm sans échelonnage. Cette ligne contient deux énoncés de sens opposé à chaque extrémité tels que « pas du tout d'accord » et « tout à fait d'accord ». Les participants devaient placer une barre verticale à l'endroit de la ligne qui correspondait le mieux à leur perception subjective. Pour l'encodage l'emplacement de la barre est

mesuré avec une réglette. Dans notre évaluation, nous avons encodé les résultats en centimètres mais pour plus de précision il est possible de mesurer les résultats en millimètre.

L'enquête par questionnaire a été complétée par des *Focus groups* avec les membres des équipes pédagogiques, dont les thèmes étaient l'apprentissage et l'évaluation des élèves.

Pour les élèves et les parents, des questionnaires papier ont été distribués, en classe pour les élèves et lors de la remise des bulletins pour les parents. L'équipe éducative a répondu à un questionnaire en ligne via une page web avec un code d'accès. L'anonymat des participants a été respecté.

3. Les résultats

Dans ce chapitre, des résultats issus des comparaisons entre les groupes (équipe éducative, élèves et parents) seront présentés. Les items choisis portent sur le tutorat individuel et hebdomadaire mis en place au lycée ainsi que la compréhension des critères d'évaluation (des performances) par les parents et les élèves eux-mêmes.

Ont répondu aux questionnaires:

- les parents des élèves du cycle inférieur (n = 136) ;
- l'équipe éducative, c'est-à-dire les enseignants (n = 46), les éducateurs gradués (n = 12) et les artisans et artistes accompagnant les activités spécialisées (n = 14) ;
- les élèves du cycle inférieur (n = 357).

L'item « *Le tutorat donne à l'élève une occasion régulière d'autoévaluation* » donne une moyenne (\pm ESM) de:

- 4.98 (\pm 0.17) pour les élèves ;
- 7.58 (\pm 0.19) pour les parents ;
- 8.10 (\pm 0.30) pour l'équipe éducative.

Nous remarquons une différence statistique significative ($F_{(2, 564)} = 54.36$; $p < 0.001$) entre les groupes (voir figure 1). Une analyse post hoc a montré que les élèves se distinguent significativement de l'opinion des parents ainsi que de l'équipe éducative. Les élèves ne considèrent pas que le tutorat leur donne une occasion régulière d'autoévaluation, alors que l'équipe éducative et les parents pensent que oui.

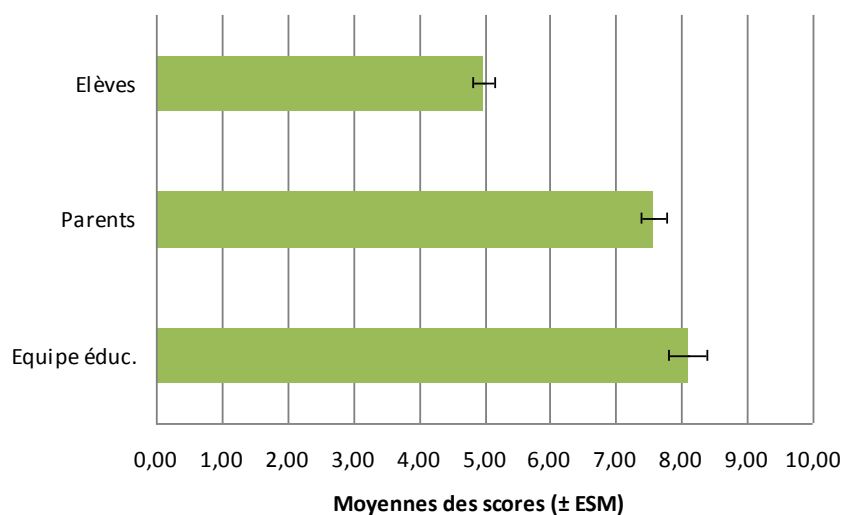


Figure 1 : Moyennes des réponses (± ESM) fournies par les différents groupes à l'item : Le tutorat donne une occasion régulière d'autoévaluation à l'élève.

Pour l'item *Le tutorat offre un accompagnement régulier à l'élève en ce qui concerne son orientation* la moyenne des scores (± ESM) est de:

- 6.28 (± 0.16) pour les élèves ;
- 7.76 (± 0.18) pour les parents ;
- 8.57 (± 0.29) pour l'équipe éducative.

De nouveau nous constatons une différence significative ($F_{(2, 513)} = 24.8$; $p < 0.001$) entre les groupes (voir figure 2) et l'analyse post hoc confirme que les élèves se distinguent significativement de l'opinion des parents ainsi que de l'équipe éducative. Les parents, mais surtout l'équipe éducative surestiment l'accompagnement offert au sein du tutorat concernant l'orientation scolaire de l'élève.

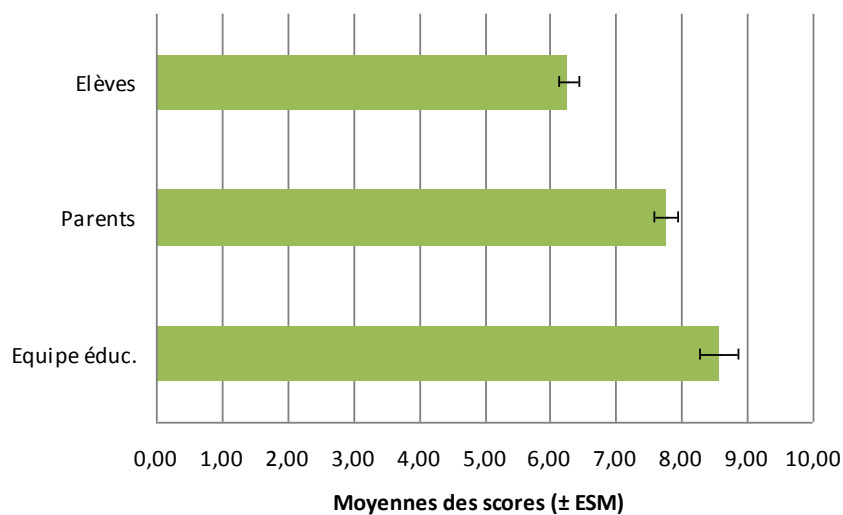


Figure 2 : Moyennes des réponses (± ESM) fournies par les différents groupes à l'item: Le tutorat offre un accompagnement régulier à l'élève en ce qui concerne son orientation.

A la question *Etes-vous satisfait du tutorat tel que pratiqué au NL?* Nous trouvons une moyenne des scores (\pm ESM) de:

- 5.85 (\pm 0.17) pour les élèves ;
- 7.96 (\pm 0.19) pour les parents ;
- 8.42 (\pm 0.29) pour l'équipe éducative.

Comme pour les autres items présentés, nous remarquons une différence statistique significative ($F_{(2, 519)} = 35.5$; $p < 0.001$) entre les groupes (voir figure 3). La perception des élèves diffère significativement de celles des membres de l'équipe éducative et des parents. Les élèves sont nettement moins satisfaits du tutorat que les parents et les membres de l'équipe éducative.

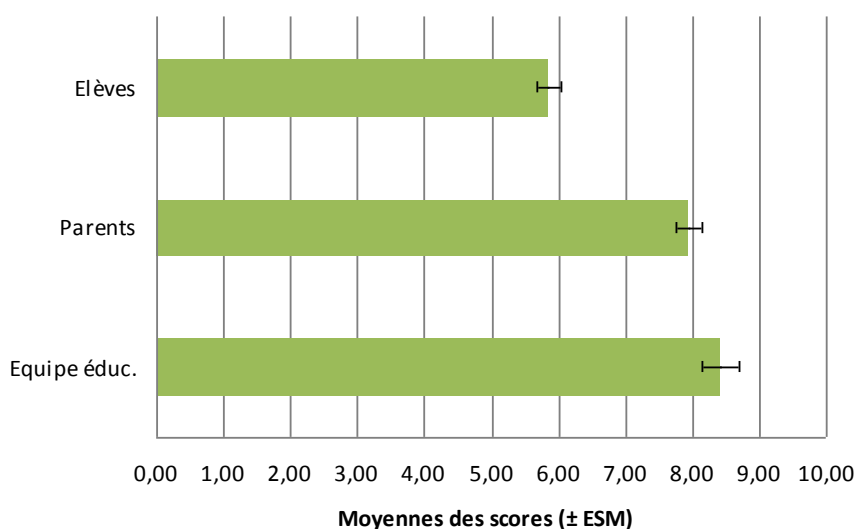


Figure 3 : Moyennes des réponses (\pm ESM) fournies par les différents groupes à l'item : *Etes-vous satisfait du tutorat tel que pratiqué au Neie Lycée ?*

Sur base de ces observations nous pouvons conclure que les élèves qui sont les premiers concernés par le tutorat, n'en sont pas très satisfaits, alors que les membres de l'équipe pédagogique et les parents pensent que le tutorat fonctionne bien.

Il a été demandé aux élèves, parents et les membres de l'équipe éducative de donner *leur perception concernant la compréhension des critères d'évaluation par les parents.*

La moyenne des scores (\pm ESM) est de:

- 6.64 (\pm 0.15) pour les élèves ;
- 7.70 (\pm 0.18) pour les parents ;
- 4.64 (\pm 0.20) pour l'équipe éducative.

Nous constatons une différence statistique significative ($F_{(2, 564)} = 42.7$; $p < 0.001$) entre ces trois groupes (voir figure 4). Dans ce cas de figure, la référence concernant ces critères d'évaluation est la perception de l'équipe éducative qui communique et applique ces-mêmes critères. Elle ne croit pas que les parents comprennent les critères appliqués pour évaluer les élèves.

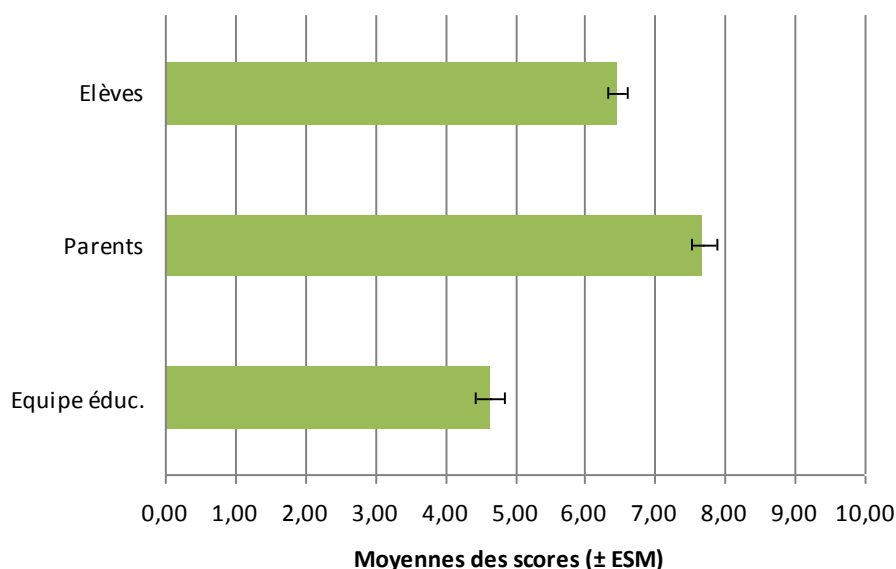


Figure 4 : Moyennes des réponses (± ESM) fournies par les différents groupes à l'item : La compréhension des critères d'évaluation par les parents.

En ce qui concerne *la compréhension des critères d'évaluation par les élèves*, nous trouvons une moyenne des scores (± ESM) (voir figure 5) de:

- 6.81 (± 0.13) pour les élèves ;
- 7.37 (± 0.2) pour les parents ;
- 5.29 (± 0.24) pour l'équipe éducative.

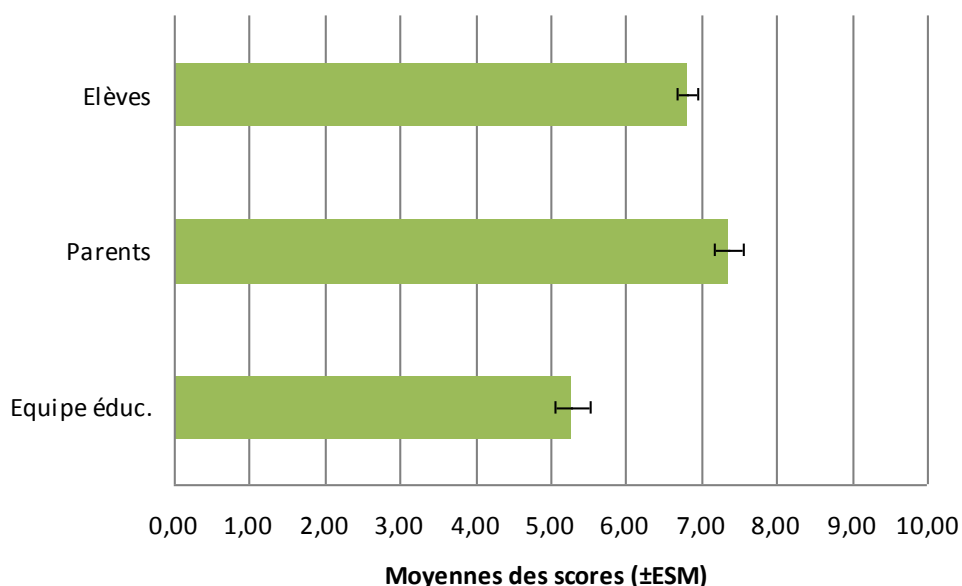


Figure 5 : Moyennes des réponses (± ESM) fournies par les différents groupes à l'item : La compréhension des critères d'évaluation par les élèves.

Nous trouvons une différence significative ($F_{(2, 565)} = 21.7 ; p < 0.001$) entre les groupes. L'équipe éducative croit que ni les élèves ni les parents ne comprennent bien les critères en vigueur pour évaluer les performances des élèves. Comme pour l'item précédent, la référence est la perception de l'équipe éducative.

4. Conclusions

La méthodologie choisie, notamment la construction de questionnaires en miroir et la création d'items sous forme d'EVA, a permis:

- de faire participer tous les groupes de la communauté scolaire à l'évaluation ;
- de croiser les points de vue des différents groupes de la communauté scolaire et de déceler les différences de perception sur l'objet selon l'appartenance à un groupe.

L'interprétation des résultats a été facilitée par le fait que dans la phase exploratoire de la recherche, les chercheurs-évaluateurs ont réalisé des observations sur le terrain et ont mené des entretiens individuels. Les *Focus groups* portant sur l'apprentissage et l'évaluation des élèves ont été complémentaires aux questionnaires.

Dans les recherches à venir, il serait intéressant de combiner l'EVA recueillant des données métriques avec des questions ouvertes fournissant des données qualitatives comme Poncelet et ses collègues (accepté) l'ont fait. Dans leur design de recherche, l'emploi de méthodes quantitative et qualitative était « complémentaire ». Ils se sont servis de la mesure quantitative comme base des entretiens directifs. Ainsi, les parents de leur échantillon ont été interviewés en parallèle de la collecte de données quantitatives sous forme d'échelles visuelles analogues (EVA). Pour chaque item, ils leur demandaient de justifier leur réponse afin d'apporter des éléments qualitatifs sur les réponses quantitatives renseignées. L'utilisation de méthodes mixtes leurs a permis d'affiner leur recherche quantitative en lui additionnant un regard qualitatif.

Le rapport final d'évaluation met en évidence les différences de perception entre les groupes et contient des informations générales quant aux principes de l'approche par compétences, l'évaluation formative et critériée ainsi que la différenciation pédagogique. Il peut servir comme point de départ pour discuter des pratiques d'apprentissage, d'évaluation et d'autoévaluation au sein de la communauté scolaire et de trouver des définitions et pratiques partagées.

En nous référant aux résultats présentés, nous concluons que dans le cadre d'une évaluation de projet ou d'établissement scolaire il est important de recueillir les perceptions de tous les acteurs concernés pour éviter des perspectives unilatérales sur un objet qui peuvent induire des interventions non pertinentes.

5. Références et bibliographie

- A'Campo, L. E. I., Spliethoff-Kamminga, N. G. A. & Macht, M. (2010). Caregiver education in Parkinson's disease: formative evaluation of a standardized program in seven European countries. *Quality of Life Research*, 19, 55-64.
- Bebell, D., O'Dwyer, L. M., Russell, M. & Hoffmann, T. (2010). Concerns, considerations and new ideas for data collection and research in educational technology studies, *Journal of Research in Educational technology studies*, 43 (1), 29-52.
- Dierendonck, C. (2010). Convergences et divergences entre parents et enfants en tant que sources d'information sur la famille, *Mesure et évaluation en éducation*, 33(2), 65-94.
- Foley, D. K. (2008). Development of a visual analogue scale to measure curriculum outcomes, *Journal of Nursing Education*, 47 (5), 209-212.
- Grant S., Aitchison T., Henderson E., Christie J., Zare S., McMurray J., and Dargie H. (1999). A comparison of the reproducibility and the sensitivity to change of visual analogue scales, borg scales, and likert scales in normal subjects during submaximal exercise. *Chest*, 116 (5): 1208-1217
- Hasson, D., Arnetz, B. B. (2005). Validation and findings comparing VAS vs. Likert scales for psychosocial Measurements, *International Electronic Journal of Health Education*, 8, 178-192.
- Myford, C. M. (2002). Investigating design features of descriptive graphic rating scales, *Applied Measurement in Education*, 15 (2), 187-215.

Actes du 25^{ème} colloque de l'ADMEE-Europe Fribourg 2013 :
Evaluation et autoévaluation, quels espaces de formation

- Poncelet, D., Pelt, V., Mancuso, G., Dierendonck, C. & Kerger, S. (accepté). Evaluation d'une recherche-action : l'apport des méthodes mixtes. *Formation et pratiques d'enseignement en questions. La revue des Hautes écoles pédagogiques et institutions assimilées de Suisse romande et du Tessin.*
- Quilter, Sh. M.; Band, J. P.; Miller, G. M. (1999). Measuring mental imagery with visual analogue scales. *Journal of Mental Health Counseling*, 21 (2) 161-72
- Turner, N. M., Leemput van de, A. J., Draaisma, J. M. T., Oosterveld, P. & Cate ten, O. Th. J. (2008). Validity of the visual analogue scale as an instrument to measure self-efficacy in resuscitation skills, *Medical Education*, 42, 503-511.