



Thématique

5. Recherche : L'étude de la validité des dispositifs d'évaluation et de leur contenu

Titre de la communication

Analyse de validité psycho-didactique d'épreuves d'évaluations internes en classe de Cours Préparatoire

Mots clés

Évaluation interne, analyse de tâches, validité psycho-didactique

Résumé court

Notre communication s'inscrit dans une réflexion sur une évaluation qui soit « au service des apprentissages ». La question à laquelle nous cherchons à répondre est celle de la validité de l'évaluation de ce point de vue : le diagnostic que l'évaluation permet est-il bien-fondé ? D'un point de vue didactique, l'évaluation pose en effet la question de sa propre validité. Nous nous attacherons dans cette communication à porter notre attention sur les tâches pour évaluer dans le cadre d'évaluations internes conduites en classe de CP (1^{ière} primaire), dans le domaine des programmes intitulé « Nombres et calcul ». Particulièrement, nous nous intéresserons à leur validité psycho-didactique. Rechercher à établir la validité psychodidactique d'une tâche revient à estimer si la tâche permet bien d'engager (de par les contenus mis en jeu, de par leur mise en contexte etc.) une activité de l'élève correspondante à ce que l'enseignant cherche à observer (ce qui est formulé par l'item).

Deux parties structureront la communication :

- dans une première partie, nous présenterons brièvement le cadre conceptuel d'analyse de la tâche que nous avons construit. Nous présenterons ensuite la méthodologie d'analyse des tâches d'évaluation que nous avons élaborée en référence au cadre précédemment défini.
 - dans une deuxième partie, nous présenterons quelques-uns des résultats de nos analyses lesquelles ont été entreprises à partir d'un corpus de 754 épreuves recueillies auprès de 16 enseignants de CP. Nous les organiserons selon certaines composantes de la tâche prescrite aux élèves susceptibles d'avoir une influence sur la validité des épreuves.
- Nous concluons notre présentation en soulignant que, au sein des épreuves que nous avons collectées et analysées, les composants mis en lumière semblent insuffisamment pris en compte. De ce fait, la qualité diagnostique de l'évaluation conduite est susceptible d'être faible.

Résumé long

Notre communication s'inscrit dans une réflexion sur une évaluation qui soit « au service des apprentissages » (expression faisant florès utilisée par exemple par le ministre de l'Éducation

nationale B. Hamon lors du lancement de la conférence nationale sur l'évaluation des élèves en juin 2014). La question à laquelle nous cherchons à répondre est celle de la validité de l'évaluation de ce point de vue : le diagnostic que l'évaluation permet est-il bien-fondé ? Au-delà du débat autour de la question de savoir si l'on évalue des compétences (De Ketele & Gérard, 2005), d'un point de vue didactique, l'évaluation pose surtout la question de sa propre validité. En 1979, Cardinet (1979, p. 216) dénonçait les évaluations qui apprécient la performance scolaire « sans prendre en compte de quoi dépend la réussite ou l'échec ». Dans une perspective éduométrique, plusieurs validités peuvent être analysées (validité de « conformité », validité de contenu (De Lansheere, 1988, p.81)) selon que l'on cherche à observer si l'évaluation correspond à la prescription d'évaluer, si les principaux composants des domaines étudiés sont couverts par l'évaluation. Nous nous attacherons dans cette

communication à porter notre attention sur les tâches pour évaluer dans le cadre d'évaluations internes conduites en classe de CP (1^{ière} primaire), dans le domaine des programmes intitulé « Nombres et calcul ». Particulièrement, nous nous intéresserons à leur validité psycho-didactique. Rechercher à établir la validité psycho-didactique d'une tâche revient à estimer si la tâche permet bien d'engager (de par les contenus mis en jeu, de par leur mise en contexte etc.) une activité de l'élève correspondante à ce que l'enseignant cherche à observer (ce qui est formulé par l'item).

Deux parties structureront la communication :

-dans une première partie, nous indiquerons que caractériser du point de vue psychodidactique la validité d'épreuves implique de considérer les tâches prises individuellement, et nécessite de s'intéresser « aux contenus impliqués dans les évaluations, mais aussi au fonctionnement cognitif des évalués, aux processus de réponse provoqués par la tâche, tâche dont les caractéristiques doivent être analysée » (Vantourout & Goasdoué ; 2014)

Pour ce faire, nous présenterons brièvement le cadre conceptuel d'analyse de la tâche que nous avons construit. Il a pour point de départ un ancrage psychologique ergonomique (Leplat 1977). Analyser une épreuve d'évaluation consiste, pour nous, à analyser une tâche prescrite sous deux angles différents mais fortement liés et indissociables : le premier considère la tâche prescrite comme un produit de l'activité de l'enseignant, un construit, le second considère la tâche prescrite comme un support de l'activité de l'élève.

Nous présenterons ensuite la méthodologie d'analyse des tâches d'évaluation que nous avons élaborée en référence au cadre précédemment défini.

Dans le cadre des approches psycho-didactiques de l'évaluation, apprécier la validité d'une tâche prescrite consiste à apprécier :

- d'une part, l'adéquation de la tâche prescrite (à l'élève) à la (double) tâche à réaliser de l'enseignant (évaluer et informer),

- d'autre part, le potentiel de la tâche prescrite à engager l'élève dans une tâche effective qui corresponde bien à ce pourquoi elle est prescrite à l'élève.

Une tâche serait réputée valide de ce point de vue si l'écart était nul entre ces différents niveaux de la même tâche. C'est-à-dire si l'épreuve que propose l'enseignant permettait bien d'évaluer l'item choisi, et si une telle tâche, dans l'idéal théorique, sous-tendait une activité de l'élève s'appuyant sur les connaissances ou compétences visées, sans que d'autres paramètres de la tâche mobilisant d'autres connaissances ou compétences n'interfèrent, et ne conduisent l'enseignant à inférer des copies des élèves des éléments erronés quant à l'activité effective de l'élève.

-dans une deuxième partie, nous présenterons quelques-uns des résultats de nos analyses lesquelles ont été entreprises à partir d'un corpus de 754 épreuves recueillies auprès de 16 enseignants de CP. Nous les organiserons selon certaines composantes de la tâche prescrite aux élèves susceptibles d'avoir une influence sur la validité des épreuves. Nous présenterons certains résultats obtenus suite à l'analyse d'épreuves de Lecture : en effet, en CP la question du savoir-lire, savoir-écrire se pose dès que l'élève est soumis à une tâche écrite, elle ne peut être éludée.

Parmi les composantes de la tâche susceptibles d'avoir une influence, nous distinguerons le dessin, susceptible d'intervenir au niveau de la représentation de la tâche ou du choix de stratégies, les questionnaires de compréhension utilisés en lecture que l'on peut considérer comme des textes en eux-mêmes voire des textes sur les textes, le système de réponse imposé ou suggéré par la tâche qui amène l'élève à produire une réponse qui sera ensuite interprétée par l'enseignant. Nous présenterons un exemple de la difficulté à opérationnaliser au moyen d'une épreuve le fait d'évaluer la capacité à résoudre un problème simple à une opération. Enfin, nous discuterons du choix des tâches particulièrement lorsqu'elles sont le fait d'un emprunt (à un site, à un matériel éditorial) ce qui a pour effet, d'une certaine manière, de considérer a priori qu'elles sont valides.

Nous concluons notre présentation en soulignant que, au sein des épreuves que nous avons collectées et analysées, les composants mis en lumière semblent insuffisamment pris en compte. De ce fait, la qualité diagnostique de l'évaluation conduite est susceptible d'être faible.