

QUELQUES RESULTATS CONCERNANT LES COMPETENCES EN RESOLUTION DE PROBLEMES D'ELEVES EVALUES SUR UN MEME PROBLEME ET A L'AIDE D'UNE MEME GRILLE D'EVALUATION

CHANUDET* Maud

Résumé – Nous présentons quelques résultats d'une étude qui s'intéresse à l'évaluation des compétences des élèves en résolution de problèmes mathématiques. Nous analysons les résultats de près de 500 élèves de 13-14 ans ayant à résoudre le même problème et étant évalués à partir des mêmes critères afin de répondre aux questions suivantes : Quels critères d'évaluation sont majoritairement validés/invalidés par les élèves ? A partir de leurs résultats, peut-on identifier différents profils d'élèves ?

Mots-clefs : évaluation, résolution de problèmes, compétences, critères d'évaluation, profil d'élèves

Abstract – We present some results of a study that focuses on students' mathematics problem solving competencies. We analyse the results of almost 500 students aged 13-14 who have to solve the same problem and who are assessed about the same grid of criteria, in order to answer to the following questions: what assessment criteria are mainly validated by students? Thanks to their results, is it possible to identify different students' profiles?

Keywords: assessment, problem solving, competencies, assessment criteria, students' profile

Nous présentons quelques résultats d'une étude qui s'intéresse à l'évaluation des compétences des élèves en résolution de problèmes. Cette étude s'inscrit dans une recherche doctorale qui questionne en particulier les pratiques évaluatives des enseignants dans le cadre d'un enseignement centré sur la résolution de problèmes.

I. CONTEXTE DE L'ETUDE

En Suisse romande, la résolution de problèmes est au cœur de l'enseignement des mathématiques et ce à tous les niveaux scolaires, que ce soit comme un *moyen* ou comme un *objectif* d'apprentissage (CIIP, 2010). Pour mettre en avant cette dimension importante de l'activité mathématique, a été créé dans le canton de Genève où est menée l'étude que nous présentons, un cours intitulé « démarches mathématiques et scientifiques » (DMS). Dispensé à raison d'une période de 45 minutes par semaine à des élèves de 13-14 ans, ce cours est centré exclusivement sur la résolution de problèmes mathématiques. Il donne par ailleurs lieu à une évaluation certificative indépendante du cours de mathématiques ordinaire et les directives officielles font reposer cette évaluation sur le dispositif de la narration de recherche (Bonafé et al., 2002). On sait cependant qu'évaluer les compétences des élèves en résolution de problèmes est complexe.

Ainsi, afin d'homogénéiser à la fois entre enseignants et entre cycles d'orientation¹ les attendus, et de proposer un outil aux enseignants pour évaluer leurs élèves, nous avons travaillé en 2015-2016 avec deux enseignantes du cours de DMS au sein d'une commission chargée de développer une grille d'évaluation des narrations de recherche dans le contexte de ce cours. Nous nous sommes pour cela appuyées sur une grille existante, issue des travaux de l'équipe genevoise impliquée dans le projet européen PRIMAS² et avons fait évoluer cette grille en fonction des retours des enseignantes membres de la commission qui l'ont utilisé pendant une année dans leur classe, et de certains résultats de recherche (caractère scientifique

* Université de Genève – Suisse - maud.chanudet@unige.ch

¹ Etablissements accueillant les élèves de 12 à 15 ans. Equivalents des collèges en France.

² Promoting inquiry-based learning in mathematics and science at both primary and secondary levels across Europe. Site du projet disponible à : <http://www.primas-project.eu>

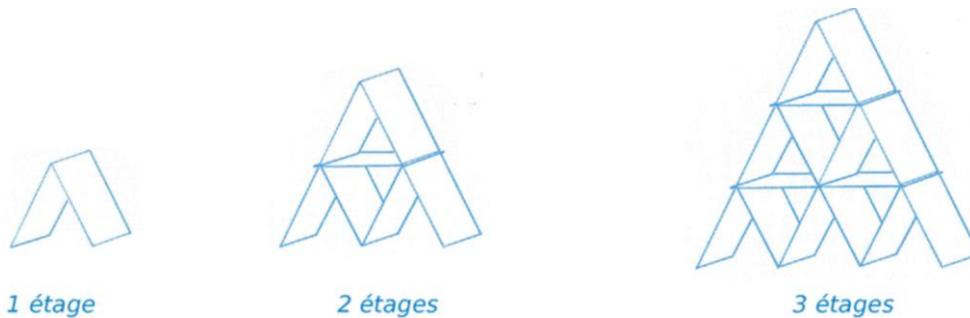
d'une démarche de type essais-conjecture-test-preuve reposant sur l'articulation entre les phases, importance de la phase d'essais et de sa dialectique avec les autres phases par exemple (Hersant, 2010)). La grille à laquelle la commission a abouti (tableau 1) est organisée autour de cinq dimensions (présentation, narration, modélisation, recherche et technique) déclinées en différents critères (texte compréhensible, narration pertinente et narration complète pour la dimension narration par exemple). Elle vise à donner aux enseignants des orientations communes quant aux points de vue à adopter sur les productions des élèves.

Dimensions	Critères
Présentation	La présentation est soignée.
Narration	Le texte est compréhensible. La narration est pertinente (centrée sur la résolution du problème). Chaque étape de la recherche est décrite de façon structurée (elle est introduite, développée et conclue). La narration est complète (toutes les étapes sont présentes). Les étapes sont présentées dans l'ordre chronologique.
Modélisation	L'élève s'est approprié le problème : reformulation, traduction en langage mathématique, schématisation, etc. Les outils, concepts mathématiques et stratégies utilisés sont pertinents.
Recherche	Une méthode de résolution est mise en œuvre, des pistes de résolution sont dégagées. Les essais sont cohérents avec le problème et visent à faire ressortir les régularités (commencent par des cas simples, sont suffisamment nombreux et variés, etc.). L'absence d'essais n'est pas pénalisée, si une conjecture valide a été trouvée et testée. Les éléments qui ont permis d'énoncer chaque conjecture sont identifiables (explications, texte souligné, code couleur, etc.) ; en particulier, s'il y en a, le lien avec les essais est exprimé. Une conjecture valide ou un nombre suffisant de conjectures non valides, est énoncé. Toute conjecture est testée et la démarche aboutit, soit à une preuve (qui la valide définitivement), soit à un contre-exemple (qui l'invalidé), soit à des tests suffisamment nombreux et variés (qui ne permettent pas de l'invalidé). Chaque conjecture fait l'objet d'une conclusion cohérente avec la démarche décrite au point précédent. A l'issue de ses recherches, l'élève exprime une conclusion (solution au problème, réponses partielles au problème, etc.) cohérente avec le problème et les recherches qu'il/elle a effectuées.
Technique	Les outils et concepts mathématiques sont utilisés correctement. Les codes, notations, symboles, qui ne font pas partie du problème, sont définis.

Tableau 1 – Grille d'évaluation des narrations élaborée par la commission.

Afin de diffuser cet outil auprès des enseignants, nous avons co-animé une formation continue obligatoire à laquelle a participé la quasi-totalité des enseignants dispensant ce cours cette année-là. Cette formation s'est déroulée sur deux demi-journées, l'une en novembre 2016, l'autre en mars 2017. La première demi-journée a été consacrée à la présentation de la grille d'évaluation et à une première prise en main de cet outil. Entre les deux demi-journées, les enseignants ont proposé dans leur classe un même problème et ont utilisé la grille d'évaluation présentée ci-dessus pour évaluer leurs élèves sur ce problème.

Le problème est celui dit « des châteaux de cartes » dont l'énoncé est le suivant³ : *Pour construire un château de cartes à un étage, il faut 2 cartes. Pour un château de cartes à deux étages, il faut 7 cartes. Et pour un château de cartes à trois étages, il faut 15 cartes. Combien faut-il de cartes pour construire un château à 7 étages ? A 30 étages ? A 100 étages ?*



II. METHODOLOGIE

Nous nous appuyons sur les données recueillies dans le cadre de cette formation pour mener une étude à grande échelle concernant les compétences des élèves en résolution de problèmes. La nécessité de pouvoir analyser un grand nombre de données (les productions de 494 élèves) nous a amené à définir un dispositif très contrôlé (un problème et des critères d'évaluation communs), cohérent avec les contraintes institutionnelles, afin de nous permettre de dégager des tendances générales.

Nous avons recueilli les résultats des élèves par le biais d'un sondage en ligne⁴. Chaque enseignant a ainsi évalué les productions de ses élèves en se prononçant sur la validation ou l'invalidation de chacun des critères de la grille. Il s'agissait dans le sondage de cocher la case correspondant au critère lorsque l'enseignant le considérait comme validé, et de ne pas la cocher lorsqu'il le considérait comme non validé. Afin de rendre les données exploitables par un logiciel d'analyse statistique, nous avons ensuite traduit cela sous forme d'une valeur numérique, respectivement 1 ou 0, pour chacun des critères et chacun des élèves⁵.

Les compétences des élèves nous sont donc données à voir par le biais des évaluations que font les enseignants des narrations de recherche des élèves. Il pourrait nous être reproché, dans un souci d'objectivité, que les évaluations sur lesquelles nous nous appuyons pour analyser les compétences des élèves ne sont pas valides parce qu'établies par différents évaluateurs. Aussi aurions-nous pu faire le choix de demander aux enseignants de nous transmettre les productions brutes de leurs élèves puis procéder nous-même à l'évaluation de ces productions. Or le nombre très important de données à recueillir aurait rendu ce travail extrêmement long et fastidieux. De plus, les études docimologiques montrent que malgré l'unicité de l'évaluateur, des biais (effet de contraste entre copies, effet de l'ordre de correction, par exemple (Leclercq, Nicaise, & Demeuse, 2013)) apparaissent lors de l'évaluation, qui plus est lorsque le nombre de productions à évaluer est aussi conséquent. Enfin, il nous semble légitime de nous appuyer sur les évaluations produites par les enseignants ne serait-ce que parce qu'elles constituent le principal moyen d'accès aux apprentissages des élèves. C'est par elles que sont reconnus institutionnellement les apprentissages des élèves.

³ Ce problème est proposé dans le manuel de Sesamath 3^e et disponible à l'adresse suivante : http://mep-outils.sesamath.net/manuel_numerique/diapo.php?atome=48667&ordre=1

⁴ Le sondage est hébergé sur le site de Limesurvey, <https://www.limesurvey.org/fr/>

⁵ Nous analysons les données de ce sondage à l'aide du logiciel d'analyse statistique SPSS.

III. QUESTIONS DE RECHERCHE

Nous nous intéressons d'une part aux résultats des élèves relativement au référentiel d'évaluation commun et au problème proposé, d'autre part aux potentiels profils d'élèves que ces résultats peuvent nous permettre d'identifier. Les questions auxquelles nous souhaitons répondre ici sont plus précisément :

- Quels critères d'évaluation les enseignants du cours de DMS valident/invalident-ils majoritairement relativement aux productions des élèves sur le problème dit des châteaux de cartes et ce, en milieu d'année scolaire ?
- A partir de leurs résultats, peut-on identifier différents profils d'élèves ?

IV. RESULTATS ET INTERPRETATION

1. Compétences des élèves

En considérant 0 et 1 comme une note attribuée par l'enseignant à la production de l'élève relativement à un critère donné, et en calculant la moyenne pour ce critère, nous obtenons à la fois une note moyenne, mais aussi la proportion d'élèves pour lesquels le critère a été considéré comme validé. Le pourcentage des productions d'élèves pour lesquelles un critère a été validé est récapitulé dans le tableau 2.

Dimension	Critères	Taux
Présentation	La présentation est soignée.	81%
Narration	Le texte est compréhensible.	80%
	La narration est pertinente. Chaque étape de la recherche est décrite de façon structurée.	73%
	La narration est complète. Les étapes sont présentées dans l'ordre chronologique.	54%
Modélisation	L'élève s'est approprié le problème : reformulation, traduction en langue mathématique, schématisation, etc.	88%
	Les outils, concepts mathématiques et stratégies utilisés sont pertinents.	81%
Recherche	Une méthode de résolution est mise en œuvre, des pistes de résolution sont dégagées.	91%
	Les essais sont cohérents avec le problème et visent à faire ressortir les régularités. L'absence d'essais n'est pas pénalisée, si une conjecture valide a été trouvée et testée.	73%
	Les éléments qui ont permis d'énoncer chaque conjecture sont identifiables ; en particulier, s'il y en a, le lien avec les essais est exprimé.	51%
	Une conjecture valide ou un nombre suffisant de conjectures non valides, est énoncé.	57%
	Toute conjecture est testée et la démarche aboutit, soit à une preuve, soit à un contre-exemple, soit à des tests suffisamment nombreux et variés.	28%
	Chaque conjecture fait l'objet d'une conclusion cohérente avec la démarche décrite au point précédent.	43%
	A l'issue de ses recherches, l'élève exprime une conclusion cohérente avec le problème et les recherches qu'il/elle a effectuées.	64%
Technique	Les outils et concepts mathématiques sont utilisés correctement.	71%
	Les codes, notations, symboles utilisés, qui ne font pas partie du problème, sont définis.	83%

Tableau 2 - Tableau de synthèse des moyennes de validation des critères dans les productions d'élèves.

Nous voyons que certains critères sont validés pour une large majorité de productions (plus de 70% des élèves) : tout d'abord ceux en lien avec les dimensions « modélisation » et « technique », mais aussi la mise en œuvre d'une méthode de résolution, la recherche de régularités à partir des essais effectués, le soin, la pertinence de la narration et le fait que les narrations des élèves soient compréhensibles.

A l'inverse, un seul critère est validé pour moins de 30% des productions d'élèves. Il s'agit du critère relatif à la nécessité de tester, prouver chaque conjecture émise. Nous avons noté lors de l'analyse a priori de cette activité que la preuve validant la formule obtenue par les élèves était difficilement accessible à ce niveau scolaire. L'énoncé complet de ce critère met cependant bien en avant le fait que lorsque la preuve n'est pas accessible pour les élèves, ces derniers peuvent tester leur conjecture sur des exemples pour montrer qu'elle n'est pas invalidée par ces exemples, ce qui était une démarche attendue dans ce problème. Le faible pourcentage de productions d'élèves pour lesquels ce critère a été validé ne s'explique donc pas seulement par le choix du problème proposé mais illustre selon nous les difficultés rencontrées par un grand nombre d'élèves.

Tous les autres critères (les éléments qui ont permis d'énoncer chaque conjecture sont identifiables ; une conjecture valide ou un nombre suffisant de conjectures non valides est énoncé ; chaque conjecture fait l'objet d'une conclusion cohérente avec la démarche décrite au point précédent ; à l'issue de ses recherches, l'élève exprime une conclusion cohérente avec le problème et les recherches qu'il/elle a effectuées ; ainsi que le caractère complet et chronologique de la narration) sont validés pour environ la moitié des productions d'élèves. Cela illustre le fait qu'un nombre important d'élèves ne maîtrisent pas encore ces aspects de la recherche et de la narration associée à cette recherche.

En synthèse, sur le problème des châteaux de cartes, les élèves sont peu nombreux à avoir des difficultés pour s'appropriier le problème ou se lancer dans la recherche, ce qui était prévisible au vu du problème proposé. Les élèves peuvent en effet facilement se lancer dans la recherche en comptant sur les schémas le nombre de cartes nécessaires, pour de faibles nombres d'étages. Sur l'aspect narratif, ils produisent majoritairement des textes compréhensibles et pertinents mais qui ne relatent pas toujours de manière complète et chronologique leur recherche. Ils mobilisent par ailleurs en général des concepts et des outils mathématiques pertinents par rapport au problème et les utilisent correctement. Concernant la dimension « recherche », ils font des essais pertinents et cherchent à en faire ressortir des régularités. Néanmoins, les éléments en lien avec la notion de conjecture mettent en difficulté un plus grand nombre d'élèves. Les conjectures font peu souvent l'objet d'une recherche de validation ou d'invalidation, elles ne sont pas toujours mises en lien avec ce qui a permis d'aboutir à celles-ci, elles ne sont pas toujours en nombre suffisant et ne font souvent pas l'objet d'une conclusion. Ce sont donc le caractère complet et chronologique de la narration ainsi que la dimension « recherche » qui posent le plus de difficultés aux élèves.

Ces analyses nous donnent des résultats globaux concernant les critères majoritairement validés et invalidés par les enseignants dans les productions d'élèves. Nous en venons à nous demander s'il est possible de caractériser plus finement les productions des élèves et ainsi déterminer des profils d'élèves.

2. Profils d'élèves

Nous cherchons à déterminer l'existence de profils d'élèves relativement à la résolution de ce problème. Nous recourrons pour cela à une analyse dite *en clusters*. Cette classification d'individus en un certain nombre de sous-groupes homogènes permet de faire apparaître une potentielle structure dans des données. Nous travaillons avec un regroupement des élèves en

fonction de leurs résultats sur ce problème, sur chacun des quinze critères de la grille. Chaque élève est donc représenté par un vecteur de dimension 15. La distance entre les élèves est estimée en fonction du carré de la distance euclidienne les séparant. L'analyse en clusters vise à regrouper les élèves en sous-groupes de proximité. Suite à de premières analyses, nous choisissons de travailler avec un regroupement en neuf clusters. Nous souhaitons en effet avoir un nombre de clusters suffisant pour identifier différents profils d'élèves, mais pas trop important pour ne pas réduire le nombre d'individus constituant chaque groupe à des valeurs négligeables, et ainsi rendre trop peu représentatifs les profils identifiés.

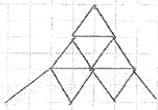
Nous choisissons ainsi une analyse en clusters qui regroupe les 494 élèves évalués en neuf sous-groupes homogènes. Trois sous-groupes sont constitués de moins de 10 individus. Nous décidons de ne pas les conserver par souci de représentativité. Les sous-groupes restants ne sont cependant pas non plus homogènes en termes de nombre d'individus concernés. En effet, tandis qu'un sous-groupe est constitué de 377 individus, les autres ne sont pas formés de plus de 41 individus. Pour caractériser chaque sous-groupe, nous considérons la moyenne obtenue à chacun des critères de la grille. Nous analysons et comparons ces moyennes à celles obtenues en considérant l'ensemble des élèves.

Le premier sous-groupe est constitué de 377 élèves. Avec 76.3% des élèves évalués c'est le sous-groupe le plus important. Ce sont des élèves qui suivent les tendances générales observées précédemment. Les moyennes des élèves de ce sous-groupe sont légèrement supérieures à celles obtenues en considérant tous les élèves, et ce pour chaque critère. Ils réussissent globalement bien et seuls quelques critères les mettent en échec. Leurs difficultés se concentrent sur la dimension « recherche » et sur le fait qu'ils n'expliquent pas toujours en détails comment ils en sont arrivés à formuler une conjecture, qu'ils ne tentent pas assez de tester et prouver leurs conjectures, et qu'ils ne concluent pas chaque étape liée à une conjecture. Nous donnons ci-dessous un extrait de production d'un élève ayant ce profil (figure 1).

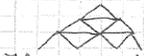
Les élèves constituant le deuxième sous-groupe sont au nombre de 41 (8.3% des élèves évalués). Ces élèves rendent un travail qui n'est pas soigné. Ils éprouvent des difficultés concernant l'aspect « narration » : leurs narrations ne sont ni pertinentes et ni complètes, et pas toujours compréhensibles. Ils éprouvent par ailleurs des difficultés quant à la dimension « recherche » : sur le fait de trouver une conjecture valide ou d'en proposer un nombre suffisant, de chercher à prouver ces conjectures et de conclure (à propos de leurs conjectures et de leur recherche). Ils maîtrisent cependant très bien la dimension « technique ». Ce sont donc principalement les compétences en lien avec la dimension « narration » et les éléments de recherche et de validation de conjectures qui leur posent problème.

Le troisième sous-groupe réunit 24 élèves (4.9% des élèves évalués). Ces élèves produisent des narrations pertinentes et compréhensibles mais qui ne sont pas complètes. Ils ont quelques difficultés à mobiliser des outils pertinents par rapport au problème. Toute la partie relative à une démarche scientifique de la dimension « recherche » leur pose problème (les phases d'essais, de conjecture et de preuve). Ils présentent enfin des difficultés importantes quant à la dimension « technique ». Ces élèves ont de bonnes compétences sur la forme de la narration, mais présentent par ailleurs des difficultés dans presque toutes les dimensions évaluées et en particulier dans les dimensions « recherche » et « technique ».

Alors, pour commencer on va essayer avec 3 étages:



On va calculer le premier étage avec une formule simple: le nombre d'étage $\cdot 2$ exemple:



Il y a 3 étages donc $3 \cdot 2 = 6$.
En somme pour le 1 étage il suffit de faire cette formule pour trouver le nombre de cartes

Mais pourquoi ($\cdot 2$) car le premier étage n'a que 2 cartes par piliers. Maintenant, il faut calculer le reste du chateau, nous allons procéder de la manière suivante:



alors on fait $(1+2) \cdot 3$
car il y a 2 piliers de 3 et 1 pilier de 3, donc 2 piliers + 1 pilier fait 3 piliers
tout d'abord, on calcule le premier étage, $3 \cdot 2 = 6$.
Maintenant, on va additionner le reste des étages, plus précisément on calcule le nombre de piliers de 3 cartes qu'il y a par étage et on fait $\cdot 3$ sans compte le 1 étage bien sûr.

Je vais encore expliquer, plus clairement:



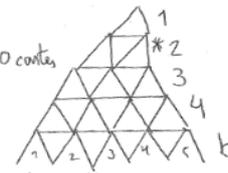
on va d'abord appeler le 1 étage: base. On va laisser la base de coté, il reste le 2 et 3. donc il suffit d'additionner ces 2 là puis faire $\cdot 3$ car il y a 3 cartes par piliers. Au début, je pensais qu'il fallait tout additionner et faire $\cdot 3$, mais j'ai remarqué que cela ne marche pas à cause de la base qui n'a que 2 cartes par piliers

Maintenant essayons cette formule avec d'autres chateaux

2 étages = 7 cartes $2 \cdot 2 = 4$

il y a un étage donc $1 \cdot 3 = 3$
pour finir $3 + 4 = 7$ cartes

5 étages = 40 cartes



nombre d'étage $\cdot 2 = 5 \cdot 2 = 10$. Donc pour la base c'est 10.
Ensuite $1 + 2 + 3 + 4 = 10$
 $10 \cdot 3 = 30$ cartes
et pour finir $30 + 10 = 40$.

Figure 1 – Extrait d'une production d'élève présentant le profil 1

Le quatrième sous-groupe est constitué de 19 élèves (3.8% des élèves évalués). Ces élèves ont des difficultés concernant la dimension « narration », sur la mobilisation d'outils mathématiques pertinents par rapport au problème, sur la dimension « recherche » et en particulier tout ce qui touche au triplet essais-conjecture-test-preuve. Ils pensent cependant davantage que la moyenne à conclure leur recherche. Ces élèves présentent donc des difficultés importantes sur la dimension « narration » et sur l'adoption d'une démarche scientifique.

Les élèves constituant le cinquième sous-groupe sont au nombre de 12 (2.4% des élèves évalués). Ces élèves rédigent des narrations pertinentes et compréhensibles mais non complètes. Ils ont des difficultés à mobiliser des outils pertinents par rapport au problème. Ils se lancent cependant sur des pistes de recherche et font des essais cohérents. Ils arrivent à trouver une conjecture valide ou ils en testent un nombre suffisant mais ne donnent pas d'explication de ce qui leur a permis d'énoncer leur(s) conjecture(s), ne cherchent pas à la/les prouver, et ne concluent pas leur recherche. Ils ont enfin des difficultés quant à la partie « technique », en particulier en ce qui concerne la mise en œuvre de concepts et d'outils mathématiques. Ce sont donc des élèves qui réussissent bien relativement à la forme de la narration et à la phase « essais » de la démarche scientifique. Ils présentent cependant des difficultés à mobiliser des outils pertinents, à justifier leur conjecture et à chercher à les prouver, tout comme quant à la maîtrise des outils et concepts mathématiques utilisés.

Le sixième et dernier sous-groupe est constitué de 11 élèves (2.2% des élèves évalués). Ces élèves présentent des difficultés importantes concernant tous les aspects de la dimension « narration ». Bien que s'étant appropriés le problème, ils ne se lancent pas sur la recherche

de pistes de résolution, ils ne mobilisent pas des outils pertinents pour tenter de résoudre le problème. Ils ont aussi des difficultés concernant le fait de trouver une conjecture valide ou d'en proposer un nombre suffisant, de chercher à les prouver et de conclure leur recherche. Ils présentent enfin des difficultés sur les deux aspects relatifs à la dimension « technique ». Ces élèves présentent donc des difficultés importantes sur toutes les dimensions évaluées.

En synthèse, le profil largement majoritaire est donc celui d'un élève qui réussit globalement bien et qui présentent des difficultés en lien avec la dimension « recherche » (le fait qu'il n'explique pas assez en détails ce qui l'a amené à formuler une conjecture, qu'il ne tente pas suffisamment de tester/prouver ses conjectures et qu'il ne conclut pas chaque étape liée à une conjecture). Pour les cinq autres profils identifiés, il nous semble intéressant de noter que tous partagent des difficultés concernant le caractère complet de la narration, la mobilisation d'outils pertinents par rapport au problème et le fait de chercher à tester/prouver les conjectures. Le seul élément validé pour tous les profils est celui relatif à l'appropriation du problème. Tous les autres critères, une majorité donc, amènent à des divergences ; ils sont maîtrisés par certains, posent des difficultés à d'autres.

V. CONCLUSIONS ET PERSPECTIVES

Cette étude à large échelle permet de mettre en lumière les compétences et les difficultés rencontrées par les élèves en lien avec la résolution de problèmes lorsque celle-ci devient un objectif d'apprentissage pour elle-même et que l'évaluation s'organise autour du dispositif de la narration de recherche. Les compétences et difficultés des élèves nous sont données à voir par le biais des évaluations que font les enseignants des productions des élèves relativement à un problème donné. Ces évaluations constituent selon nous un matériau riche pour établir, à un moment donné, un bilan des compétences d'un grand nombre d'élèves. Il nous semblerait intéressant de compléter cette étude effectuée à un niveau macro par une étude plus fine des compétences mises en œuvre par quelques élèves lors de la résolution de tels problèmes. Nous n'avons de plus présenté ici qu'une partie de la recherche qui nous a amené par ailleurs à interroger l'influence des critères pour lesquels les élèves rencontrent des difficultés sur la note qui leur est attribuée et sur leur réussite à la tâche ; à analyser de potentielles corrélations entre les quinze critères de la grille ; et à chercher à savoir si certains critères d'évaluation amènent à d'avantage de divergence d'interprétation entre enseignants que d'autres. Ces autres résultats ont été présentés lors du colloque EMF2018.

REFERENCES

- Bonafé F., Sauter M., Chevallier A., Combes M.-C., Deville A., Dray L., Robert J.-P. (2002) Les narrations de recherche du primaire au Lycée. *Brochure APMEP*, 151.
- Conférences Intercantonale de l'Instruction publique de la Suisse romande et du Tessin. (2010) Plan Etudes Roman 3e cycle, Mathématiques et sciences de la nature. Sciences humaines et sociale. Consulté à l'adresse <http://www.plandetudes.ch>
- Hersant M. (2010) *Empirisme et rationalité au cycle 3, vers la preuve en mathématiques*. Mémoire complémentaire pour l'Habilitation à diriger des recherches, Université de Nantes.
- Leclercq D., Nicaise J., Demeuse M. (2013) Docimologie critique: des difficultés de noter des copies et d'attribuer des notes aux élèves. In M. Demeuse, *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation* (pp. 273-292). Les éditions de l'Université de Liège.