



## Programme d'entraînement pour les résultats d'apprentissage : méthode d'établissement et d'évaluation des résultats d'apprentissage fondée sur les travaux scolaires

Steve Joordens, Dwayne Paré  
et Lisa-Marie Collimore  
Laboratoire de technologies pour l'enseignement  
supérieur, Université de Toronto Scarborough

Publié par

## le Conseil ontarien de la qualité de l'enseignement supérieur

1, rue Yonge, bureau 2402  
Toronto (Ontario) Canada M5E 1E5

Téléphone : 416 212-3893  
Télécopieur : 416 212-3899  
Site Web : [www.heqco.ca](http://www.heqco.ca)  
Courriel : [info@heqco.ca](mailto:info@heqco.ca)

### La présente publication doit être citée de la façon suivante :

JOORDENS, S., D. PARÉ et L.-M. COLLIMORE. *Programme d'entraînement pour les résultats d'apprentissage : méthode d'établissement et d'évaluation des résultats d'apprentissage fondée sur les travaux scolaires*, Toronto, Conseil ontarien de la qualité de l'enseignement supérieur, 2014.



Les opinions exprimées dans le présent rapport de recherche sont celles des auteurs et ne reflètent pas nécessairement le point de vue ou les politiques officielles du Conseil ontarien de la qualité de l'enseignement supérieur ou d'autres organismes qui ont offert leur soutien, financier ou autre, au projet.  
© Imprimeur de la Reine pour l'Ontario, 2014

## Sommaire

S'il est vrai que la définition des résultats d'apprentissage de base pour divers programmes d'études postsecondaires suscite un grand intérêt et a fait de grands progrès, on n'a néanmoins pas beaucoup avancé dans la détermination de moyens efficaces d'évaluer l'atteinte de ces résultats. Or, en l'absence de bonnes méthodes d'évaluation, il s'avère difficile de revoir des cours ou des programmes en se basant sur les objectifs d'apprentissage.

Jusqu'ici, les mesures visant à remédier aux lacunes dans « l'évaluation des résultats d'apprentissage » reposaient pour la plupart sur l'utilisation d'outils d'évaluation qualitative distribués aux étudiantes et étudiants à leur départ d'un établissement ou à la fin d'un programme, comme le questionnaire du National Survey of Student Engagement (NSSE) et l'indice du Collegiate Learning Assessment (CLA). Ces outils ont toutefois comme défauts d'être basés sur des rapports subjectifs, d'être souvent difficiles à administrer à grande échelle de façon régulière et de ne fournir la plupart du temps qu'une vue d'ensemble « après coup ». Dans le présent rapport, on présente et cautionne une méthode différente fondée sur les travaux scolaires ayant un potentiel beaucoup plus grand de procurer des données quantitatives beaucoup plus détaillées.

Pour illustrer cette méthode fondée sur les travaux scolaires, on peut faire l'analogie avec les exercices de renforcement musculaire pour représenter l'objectif ultime, soit permettre à la fois d'acquérir et d'évaluer les compétences durant chaque travail scolaire. Par exemple, quand on s'exerce à soulever des haltères, on développe les muscles nécessaires pour soulever les poids en plus de mesurer exactement la force des muscles en question. De même, tous les résultats d'apprentissage les plus importants s'appuient sur des compétences intellectuelles de base qui ne peuvent être acquises que si elles sont exercées régulièrement. C'est donc dire que si nous arrivons à mettre au point des outils technologiques permettant à la fois d'exercer ces compétences et d'en faire une évaluation quantitative, il serait possible de mesurer leur évolution d'un travail scolaire à l'autre, et on obtiendrait des données beaucoup plus exactes sur l'atteinte des résultats d'apprentissage associés aux compétences en question.

Le présent rapport vise avant tout à présenter un exemple concret de la manière dont une méthode fondée sur les travaux scolaires pourrait être établie et à fournir quelques données initiales montrant les avantages qu'il y a à appliquer et à évaluer les résultats d'apprentissage de cette manière. Nous commencerons par énoncer quelques objectifs d'apprentissage primordiaux sur lesquels tous les membres du personnel enseignant s'entendent, notamment le jugement critique, la créativité, l'auto-réflexion et les communications efficaces, tant en mode réceptif qu'expressif. Nous prendrons ensuite l'exemple d'un outil technologique en particulier, soit peerScholar, pour montrer que l'utilisation d'une méthode fondée sur les travaux scolaires est bel et bien réalisable, malgré les contraintes actuelles du système d'enseignement supérieur. Enfin, deux expériences, l'une illustrant le développement de l'auto-réflexion, et l'autre, le développement du jugement critique, serviront à faire ressortir le potentiel de la méthode.

Nous tirerons ensuite la conclusion que l'exemple de peerScholar pourrait être appliqué à d'autres outils technologiques d'apprentissage et, surtout si les outils en question étaient conçus de manière à faciliter l'évaluation des objectifs d'apprentissage, il nous serait possible de créer une trousse à outils nous permettant à la fois de nous faire progresser en tant que membres du personnel enseignant et de mesurer les progrès de manière efficace en plus de fournir une foule de renseignements. Toute la valeur de l'établissement d'objectifs d'apprentissage réside, jusqu'à un certain point, dans notre capacité de bien évaluer les progrès par rapport à ces objectifs et, à notre avis, les membres du personnel enseignant auraient avantage à songer à des méthodes fondées sur les travaux scolaires pour combler ce besoin.

## Table des matières

Introduction .....	4
Résultats d'apprentissage.....	5
Connaissance de la matière .....	5
Jugement critique .....	5
Créativité.....	6
Auto-réflexion.....	6
Habilités de communication.....	6
Habilités de collaboration .....	6
peerScholar.....	8
Étape de création.....	8
Étape d'évaluation .....	8
Étape de réflexion et de révision .....	10
Quantification des résultats d'apprentissage .....	12
Des données probantes .....	18
Première expérience : auto-réflexion .....	18
Personnes participantes .....	18
Travail et documents présentés.....	18
Démarche .....	18
Résultats et analyse .....	19
Deuxième expérience : jugement critique .....	19
Personnes participantes .....	20
Travail et documents présentés.....	20
Démarche .....	20
Résultats et analyse .....	20
Sommaire.....	22
Bibliographie .....	23

## Liste des tableaux

Tableau 1 : Critères d'évaluation quantitatifs à partir de peerScholar et relation avec les objectifs d'apprentissage de base	17
Tableau 2 : Scores de l'écart absolu moyen pour le jugement critique	24

## Liste des figures

Figure 1 : Étapes du processus partiel de peerScholar pour les travaux scolaires d'après les résultats d'apprentissage de base courants	13
Figure 2 : Étapes du processus complet de peerScholar pour les travaux scolaires d'après les résultats d'apprentissage de base courants	16

## Introduction

Personne n'y échappe. Il nous arrive à tous de nous regarder dans le miroir et de nous sentir insatisfaits de ce que nous y voyons. Nous voulons améliorer notre force et notre forme physique. Poussés par notre motivation, nous décidons de la manière de procéder pour atteindre nos objectifs. La première étape consiste à déterminer ce que nous voulons changer. Voulons-nous perdre du poids? Avons-nous aussi l'intention d'augmenter notre masse musculaire? Avons-nous besoin de chirurgie esthétique? L'étape de la planification est presque amusante. Nous nous imaginons dans l'avenir, de manière plus ou moins abstraite, même si nous sommes pleinement conscients de la réalité telle qu'elle est. Ce que nous imaginons ne pourra se réaliser que si nous nous appliquons vraiment à bien planifier ce qu'il faut faire pour atteindre nos objectifs et que nous prenons ensuite des moyens concrets de manière efficace.

Ces dernières années, le système d'éducation public a aussi commencé à bien s'observer et il n'est pas non plus très fier de ce qu'il voit. Avec le temps, le nombre d'étudiantes et d'étudiants par classe a augmenté et les budgets ont diminué. Les changements que cela a entraînés dans les méthodes d'enseignement ne servent pas bien les objectifs. Il faut donc commencer par déterminer à quoi nous voulons que le système ressemble, ce qui amène habituellement à définir des objectifs d'apprentissage relativement abstraits, soit des énoncés des connaissances et des compétences transférables que nous souhaitons faire acquérir aux étudiantes et étudiants sous notre tutelle. Mais encore une fois, s'il peut être amusant de se projeter dans l'avenir, pour vraiment arriver à changer les choses, il faut avoir un plan d'action concret qui soit réalisable et sensé et qui permette de mesurer les progrès. C'est donc dire qu'il faudra s'exercer pour être à la hauteur de notre vision ambitieuse des objectifs d'apprentissage.

Pour reprendre encore une fois la même analogie, on peut parvenir à une transformation physique en faisant des exercices bien précis qui permettront d'obtenir les résultats souhaités, et ce qu'il y a d'extraordinaire dans les exercices physiques, c'est qu'ils combinent ce qu'on pourrait appeler en enseignement la pratique et l'évaluation. Par exemple, si je veux beaucoup augmenter le volume de mes biceps, des flexions des avant-bras pourraient être un bon point de départ. Si je fais cet « exercice » régulièrement, mes biceps prendront effectivement du volume et j'observerai, avec le temps, que je suis capable de soulever des poids plus lourds ou de faire plus de flexions avec les mêmes poids. Autrement dit, je constate les résultats, ce qui signifie que les « exercices » permettent du même coup une « évaluation » des progrès. Les bonnes évaluations motivent à continuer à s'exercer, ce qui a un effet d'entraînement positif qui aboutit fatalement à... une transformation en Arnold Schwarzenegger? Enfin, toute bonne analogie a ses limites, mais si on remplace le mot « biceps » par « jugement critique » et « pectoraux » par « communication verbale », on pourrait dire que notre objectif est de produire des maîtres dans l'art de développer les connaissances.

L'objectif du présent document est de faire la démonstration relativement simple qu'une combinaison semblable d'exercice et d'évaluation peut être applicable au contexte de l'enseignement pour nous aider à réaliser nos objectifs d'apprentissage. De plus, pour fournir l'exemple le plus concret possible, nous nous concentrerons sur peerScholar, une application basée sur Internet qui a été conçue pour permettre d'atteindre les objectifs d'apprentissage parmi les plus souhaités (c.-à-d. le jugement critique, la créativité, l'auto-réflexion, les communications claires et efficaces, la formulation de commentaires et la réaction appropriée aux commentaires). L'outil peerScholar combine l'évaluation par les pairs et l'autoévaluation à

l'intérieur d'un processus d'apprentissage formatif, et il y a donc de très bonnes raisons de croire que cela procure aux étudiantes et étudiants l'exercice mental nécessaire sur les types d'habiletés transférables dont ils ont besoin pour devenir productifs avant la fin de leurs études. On doit aussi se demander si cela permettra d'évaluer les résultats d'apprentissage en plus de les exercer? Autrement dit, est-ce que peerScholar, et les autres outils semblables, peuvent offrir des travaux concrets qui serviront à la fois à s'exercer pour progresser et à évaluer les objectifs d'apprentissage en classe? Et, ce faisant, ces travaux fournissent-ils un exemple de ce que les établissements d'enseignement peuvent faire pour réaliser les objectifs d'apprentissage de façon claire et efficace?

## Résultats d'apprentissage

Bien des intervenants et des organisations dans le monde de l'enseignement ont déployé des efforts considérables pour définir l'ensemble des résultats d'apprentissage traduisant les compétences que les étudiantes et étudiants doivent posséder à la fin de leurs études<sup>1</sup>. Dernièrement, le premier auteur du présent document a fait partie du comité du COQES sur les résultats d'apprentissage pour le projet Tuning qui était composé d'enseignantes et d'enseignants des niveaux universitaire et collégial et a consacré plus d'une année à la définition des objectifs d'apprentissage représentatifs des compétences et des connaissances des étudiantes et étudiants obtenant un diplôme en sciences sociales. Si la liste des objectifs d'apprentissage précis variait selon le groupe, il y avait néanmoins quelques objectifs qui revenaient pratiquement dans toutes les listes, et ils sont indiqués ci-dessous.

### Connaissance de la matière

De toute évidence, peu importe la matière, l'un des objectifs premiers de tout enseignant et de toute enseignante est de transmettre les éléments clés de la matière aux étudiantes et étudiants et, avec de la chance, aussi une partie des détails. La connaissance de la matière est considérée comme quelque peu différente des autres éléments de la liste puisque, contrairement aux habiletés cognitives, la connaissance peut s'acquérir par la simple exposition. De plus, puisque les connaissances sont évaluées régulièrement au moyen de tests conventionnels, le plus difficile pour l'enseignante ou l'enseignant est d'enseigner et d'évaluer les habiletés cognitives transférables.

### Jugement critique

L'habileté transférable que nous tenons probablement le plus à faire acquérir aux étudiantes et étudiants est la capacité de réfléchir de façon critique. Le jugement critique a diverses facettes et comprend la capacité d'analyser une chose pour en évaluer la qualité, souvent en la comparant à d'autres ou encore en la décomposant en fonction d'une certaine mesure de la qualité (Foundation for Critical Thinking 2013). Le but ultime est d'évaluer la qualité, en général en prêtant attention aux aspects moins bons, moins convaincants ou moins exacts qu'ils pourraient l'être. C'est donc dire que les étudiantes et étudiants ayant un bon jugement

---

<sup>1</sup> Consultez, par exemple, le site de l'International Society for Technology in Education (<https://www.iste.org/>) ou de Partnership for 21st Century Skills (<http://www.p21.org/>).

critique sont capables d'évaluer l'information de façon intelligente, en faisant la distinction entre les aspects qui sont convaincants et ceux qu'il vaut mieux ignorer.

## Créativité

La créativité commence souvent là où se termine le jugement critique. C'est qu'une fois que quelque chose a fait l'objet d'une évaluation critique, on est souvent conscient des lacunes et on peut avoir des idées sur la façon d'apporter des améliorations. La créativité suppose donc la capacité de percevoir comment quelque chose pourrait être transformé et de se faire une idée des coûts et des avantages du résultat imaginé. Le jugement critique permet de déceler les problèmes, tandis que la créativité sert à trouver des solutions. On peut réussir dans la vie avec seulement la capacité de réfléchir de façon critique (en étant critique de cinéma, par exemple), mais les personnes qui sont capables de formuler des solutions nouvelles et efficaces pour des problèmes existants sont celles qui peuvent améliorer les choses, et nous espérons assurément que nos étudiantes et étudiants feront partie des personnes capables de changer le monde.

## Auto-réflexion

C'est ce qu'on appelle aussi parfois métacognition. Il s'agit de la capacité de s'analyser soi-même de façon critique en ayant une excellente idée de ce qu'on sait ou pas et de ce qu'on est capable de faire ou pas et en étant conscient à la fois de ses forces et de ses faiblesses. Il est pratiquement impossible pour quiconque de s'améliorer sans savoir au départ quels sont nos points forts et nos points faibles. On pourrait dire que l'auto-réflexion équivaut, dans une certaine mesure, à la capacité d'exercer un jugement critique à son propre égard et à l'égard des produits de son propre travail.

## Habiletés de communication

On pourrait presque aller jusqu'à dire que ça ne sert pas à grand-chose qu'une personne ait de bonnes habiletés de réflexion si elle est incapable de communiquer ses idées ou d'écouter et de comprendre ce que disent les autres. Ces habiletés peuvent être considérées comme un ensemble de sous-compétences, comme la communication écrite efficace, la communication verbale efficace et la bonne compréhension des communications tant écrites que verbales. On peut aussi définir ces habiletés de façon plus précise en disant qu'il s'agit de la capacité de faire des commentaires clairs et utiles aux autres et également d'évaluer les commentaires des autres sur son propre travail et d'y réagir correctement. Cependant, de façon plus générale, on peut dire que c'est la capacité de transmettre et de recevoir des idées de façon efficace.

## Habiletés de collaboration

L'être humain est une créature très sociable et, même si une personne a de grandes habiletés de réflexion et est capable de bien communiquer, il demeure très difficile pour elle d'accomplir à elle seule quoi que ce soit qui ait de la valeur. Voilà pourquoi il est si important que les étudiantes et étudiants fassent partie de groupes sociaux et travaillent en collaboration de manière à obtenir des résultats positifs pour tous.

À noter que les six résultats d'apprentissage indiqués ci-dessus reviennent régulièrement, mais la liste n'est pourtant pas exhaustive. Par exemple, nous souhaiterions aussi que les étudiantes et étudiants sachent travailler avec les données numériques et les interpréter, notamment celles qui apparaissent dans des

tableaux et des figures. Nous voudrions aussi qu'à la fin de leurs études, ils et elles aient une plus grande conscience sociale, un meilleur sens des responsabilités et de meilleures habitudes de travail qu'à leur admission dans l'établissement. Il n'en reste pas moins que la liste qui précède contient ce que d'aucuns pourraient considérer comme les objectifs d'apprentissage les plus importants et les plus difficiles à atteindre, ce qui explique pourquoi nous les mentionnons ici.

Soulignons également que, si les objectifs d'apprentissage relatif au « contenu » sont propres à la matière d'un cours donné, les autres ont traités aux habiletés cognitives générales. Il s'agit là d'une distinction importante qui a une influence sur la façon d'enseigner, et ce, pour la raison suivante : la mémoire humaine n'assimile pas l'information de la même manière que les habiletés (Tulving 1985). L'information peut être apprise au moyen d'une simple exposition, quoique, de façon générale, plus la réflexion est approfondie, meilleur est l'apprentissage (Craik et Lockhart 1972). En revanche, la seule façon d'acquérir des compétences est de bien les exercer régulièrement (Milner 1962). Voici un exemple qui illustre bien cette distinction : en suivant un cours de deux heures, vous pouvez en apprendre beaucoup sur un domaine, notamment le karaté. Par contre, pour pouvoir « faire » du karaté, c'est-à-dire acquérir les habiletés de base nécessaires, il faut s'exercer régulièrement à faire les mouvements de base de la bonne façon.

Les habiletés cognitives ne font pas exception à cette règle. Il n'est pas possible d'expliquer à quelqu'un comment réfléchir de façon critique ou bien écrire. L'étudiante ou l'étudiant doit s'exercer régulièrement, de préférence dans un contexte structuré et encadré. C'est exactement pour cette raison qu'il est si difficile d'enseigner ces habiletés avec un trop grand nombre d'étudiantes et d'étudiants par classe et avec d'aussi petits budgets. Les méthodes traditionnelles d'enseignement, qui exigent que l'enseignant ou l'enseignante prenne part à chaque étape de l'apprentissage, sont coûteuses et complexes sur le plan logistique et c'est pourquoi elles sont souvent délaissées au profit de méthodes qui facilitent la formation et l'évaluation (l'apprentissage de la matière au moyen de cours magistraux et de tests à choix multiples).

Il peut ressortir quelque chose de positif de tout cela, en quelque sorte, si on arrive à trouver un moyen efficace de donner aux étudiantes et étudiants l'occasion d'exercer leurs habiletés cognitives autant qu'il est nécessaire. Lorsque des habiletés sont suffisamment exercées, elles deviennent pratiquement automatiques et extrêmement perfectionnées (Samuels et Flor 1997). Prenons Jimi Hendrix. Il n'avait pas à penser à bien jouer sa musique; ça se faisait naturellement. Bruce Lee, quant à lui, n'avait pas à réfléchir à ses mouvements de kung-fu; chaque réaction se faisait par réflexe et, une fois apprise, restait bien ancrée pour toujours. C'est donc dire que, si nous donnons aux étudiantes et étudiants la possibilité de beaucoup s'exercer à appliquer les habiletés cognitives correspondant aux résultats d'apprentissage dont la liste a déjà été fournie, les habiletés en question pourraient aussi devenir pratiquement automatiques et ils et elles pourraient les conserver bien après la fin de leurs études.

Tout cela nous amène, bien entendu, à nous poser deux questions : Comment faut-il procéder exactement pour enseigner les habiletés cognitives à la base de ces résultats d'apprentissage dans la conjoncture économique actuelle? Est-il possible d'évaluer les objectifs d'apprentissage en cours de route de manière à pouvoir faire le suivi des progrès? Encore une fois, ce qu'il faudrait, c'est une sorte d'« entraînement cognitif » qui permettrait à la fois de s'exercer pour atteindre les résultats d'apprentissage considérés comme essentiels et d'évaluer les progrès.

## peerScholar

Maintenant que le contexte général a été présenté, le reste du rapport portera sur un outil technologique en particulier et servira à démontrer comment il peut permettre à la fois d'aider les étudiantes et étudiants à obtenir les résultats d'apprentissage visés et d'évaluer leurs progrès en cours de route. Notre choix s'est arrêté sur peerScholar, et ce, pour deux raisons. D'abord, sur le plan pratique, il s'agit d'un outil mis au point par notre laboratoire de technologies pour l'enseignement supérieur. C'est donc dire que nous avons la capacité et le code nécessaires pour explorer les données plus que nous pourrions le faire avec un outil qui aurait été conçu par quelqu'un d'autre. Ensuite, sur le plan théorique, nous ne connaissons tout simplement aucun autre outil qui puisse contribuer à faire avancer tout l'éventail des objectifs d'apprentissage en s'appuyant sur une expérience aussi approfondie que peerScholar, même que c'est expressément dans ce but que peerScholar a été conçu (Joordens, Desa et Paré 2009).

Dans le reste de la présente section, nous verrons les étapes d'un travail scolaire typique réalisé à l'aide de peerScholar, en faisant ressortir les résultats d'apprentissage soutenus au fur et à mesure. En règle générale, les travaux basés sur peerScholar amènent les étudiantes et étudiants à réaliser trois étapes séquentielles pour lesquelles on leur demande d'employer des habiletés cognitives diverses, qui se chevauchent parfois. À la première étape, l'étudiante ou l'étudiant produit un travail conformément aux directives reçues, puis il ou elle évalue un sous-ensemble de travaux présentés par des camarades et termine en examinant les commentaires de ses camarades et en apportant des modifications à son travail original en fonction de son analyse des commentaires. Il ou elle peut aussi être appelé en cours de route à évaluer la qualité de son propre travail en lui attribuant une note à l'étape que choisit l'enseignant ou l'enseignante.

### Étape de création

À la première étape d'un travail effectué à l'aide de peerScholar, l'étudiante ou l'étudiant doit soumettre un travail quelconque sur support informatique, en suivant les directives de l'enseignante ou de l'enseignant. Les résultats d'apprentissage associés à cette étape sont reliés en grande partie au travail défini par l'enseignante ou l'enseignant. L'étudiante ou l'étudiant peut avoir à faire preuve de jugement critique ou de créativité ou à faire une réflexion et il ou elle peut présenter son travail sous forme de texte, d'images ou de film (p. ex. une présentation verbale). En fin de compte, le travail de l'étudiante ou l'étudiant sera noté, ce qui signifie que l'étape de création donne à l'enseignante ou l'enseignant toute la latitude voulue pour faire porter l'exercice et l'évaluation sur l'objectif d'apprentissage de son choix. De plus, le temps consacré par l'étudiante ou l'étudiant à l'exercice du résultat d'apprentissage en question est essentiellement multiplié aux étapes suivantes.

### Étape d'évaluation

Dans le livre *Traité du zen et de l'entretien des motocyclettes* écrit en 1974, Robert Persig parle d'un travail qu'il a donné dans un cours de rhétorique. Il fallait rédiger une argumentation sans titre. M. Persig a recueilli les feuilles, les a affichées sur le mur et a demandé aux étudiantes et étudiants de les lire et de déterminer quelle était la meilleure argumentation en justifiant leur réponse. C'est que M. Persig était convaincu que la capacité de percevoir la qualité est un réflexe primitif et qu'il était donc possible pour les étudiantes et étudiants d'évaluer quelle était la meilleure argumentation, ce qui devait les amener à faire une expérience

d'apprentissage personnelle et concrète en essayant d'exprimer pourquoi une argumentation semblait meilleure qu'une autre<sup>2</sup>.

Le fonctionnement de peerScholar se rapproche beaucoup du processus d'apprentissage entre pairs décrit par M. Persig. Après avoir soumis leur travail, les étudiantes et étudiants ouvrent une session dans le système et reçoivent un sous-ensemble de travaux anonymes (disons six) de camarades choisis au hasard. Ils doivent commencer par lire tous les travaux tout en se demandant quel est celui qui est de la meilleure qualité. Ils se basent ensuite sur cette vague notion de qualité pour attribuer une note en fonction d'une échelle déterminée pour rendre les choses plus concrètes. Enfin, ils doivent présenter deux commentaires relatifs à chaque travail : un premier pour souligner quelque chose qui leur a réellement plu et un deuxième plus crucial, soit un commentaire constructif indiquant de façon claire à l'auteur ce qui, à leur avis, pourrait être changé de manière à améliorer le plus possible la qualité. En formulant leurs commentaires, les étudiantes et étudiants doivent d'abord réfléchir aux différents moyens qui pourraient être pris pour améliorer le travail, pour ensuite se concentrer sur le changement le plus important. Ils ont pour directive de ne pas se contenter de signaler un problème mais d'également donner des indications précises sur la façon de le régler de façon efficace.

Pour ce qui est de l'exercice des objectifs d'apprentissage, là aussi le processus d'évaluation du travail des pairs exige forcément un jugement critique et une auto-réflexion. En effet, les étudiantes et étudiants sont appelés à faire une analyse, des comparaisons et une évaluation pour en venir à définir un degré de qualité précis par rapport aux autres travaux auxquels ils ont accès. Ce sont là toutes les facettes du concept de jugement critique. De plus, nous exposons de cette façon les étudiantes et étudiants aux travaux réalisés par leurs pairs, ce qui les amène inévitablement à comparer leur propre travail à celui de leurs pairs et à se faire une idée bien claire et concrète de la qualité de leur travail par rapport aux autres et de ce qu'ils peuvent faire pour l'améliorer. En fait, on peut pousser encore plus loin l'auto-réflexion en demandant aux étudiantes et étudiants d'évaluer aussi leur propre travail de façon explicite, une démarche qui peut être intégrée à n'importe quelle étape.

En se basant sur leur analyse pour formuler des commentaires, les étudiantes et étudiants mettent en pratique leurs habiletés de communication et de collaboration tout en exerçant, jusqu'à un certain point, leur créativité. Il est relativement évident que les habiletés de communication sont exercées, mais c'est le ton de la communication, en particulier du commentaire constructif, qui compte vraiment. Ce qu'on demande aux étudiantes et étudiants dans le fond, c'est d'aider leurs camarades de classe à s'améliorer, ce qui est manifestement une forme de collaboration. De plus, le fait qu'ils aient à dire à leurs pairs comment faire pour s'améliorer est d'une importance primordiale puisque c'est là que la créativité entre en jeu. La reconnaissance d'un problème exige un jugement critique, sa résolution nécessite de la créativité, la formulation de la solution exige des habiletés de communication et d'exécuter le tout de manière vraiment utile illustre bien la collaboration.

Avant de clore la section sur l'étape d'évaluation, il importe de revenir sur un point qui a déjà été abordé. Tous les processus déjà décrits se déroulent dans le contexte d'un travail en particulier qui pourrait être conçu de manière à permettre d'exercer un autre objectif d'apprentissage. Par exemple, le travail pourrait

---

<sup>2</sup> L'influence de l'évaluation par les pairs sur l'apprentissage cadre bien avec l'opinion exprimée par Vygotsky (1978), et de nombreuses études viennent l'étayer tant sur le plan empirique que théorique (p. ex. King 2002; Venables et Summit 2003).

consister à faire une rédaction sur un aspect de la responsabilité sociale. Dans ce cas, l'exercice du jugement critique et des habiletés de créativité, d'auto-réflexion, de communication et de collaboration serait orienté vers la responsabilité sociale. Par conséquent, les premières réflexions individuelles sur la responsabilité sociale seraient enrichies du fait qu'il faudrait réfléchir aux idées développées par les camarades, les évaluer et les commenter.

Pour ce qui est de la logistique et des ressources, à noter que l'apprentissage approfondi dont il est question s'effectue sans l'intervention du personnel enseignant. En effet, les étudiantes et étudiants apprennent en évaluant et en commentant les travaux des pairs, et c'est le système qui gère la logistique. C'est aussi vrai pour l'étape suivante, soit l'étape de réflexion et de révision.

## Étape de réflexion et de révision

À cette étape, l'étudiante ou l'étudiant prend connaissance des notes et des commentaires fournis par ses pairs. Rappelons qu'il y a normalement six notes, six commentaires positifs et six commentaires constructifs, comme l'indique la section précédente. Il faut signaler à l'avance que, peu importe ce qu'on fait dans la vie, il y a nécessairement des pairs qui font des commentaires. Il arrive que les commentaires soient utiles, mais pas toujours, et il faut absolument que les étudiantes et étudiants apprennent à juger les commentaires qui leur sont faits. Voilà pourquoi on leur demande d'évaluer de façon explicite la valeur des commentaires reçus en leur attribuant la cote « inutile », « utile » ou « très utile ».

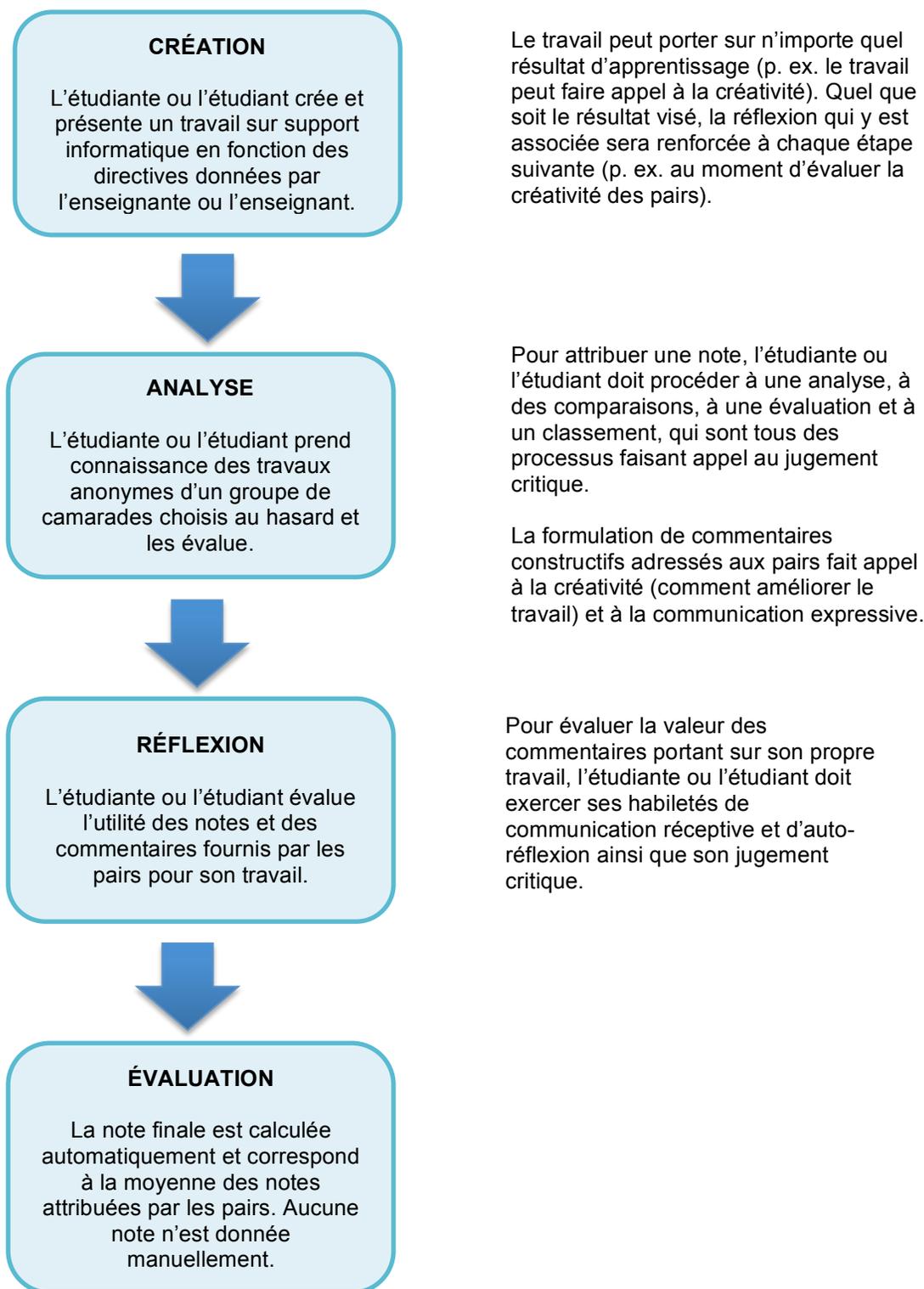
Avant d'aller plus loin, examinons le processus du point de vue des résultats d'apprentissage définis. Les étudiantes et étudiants évaluent (jugement critique) les commentaires écrits (communication réceptive) portant sur leurs propres travaux (auto-réflexion), qui, avec de la chance, sont eux-mêmes assortis de résultats d'apprentissage. On peut donc dire que bon nombre des résultats d'apprentissage exercés durant l'étape d'évaluation sont exercés encore une fois, mais dans le contexte plus personnel des propres travaux de l'étudiante ou de l'étudiant. Il y a un principe qui veut que des exercices répétés dans différents contextes permettent de généraliser l'acquisition des habiletés.

Les travaux réalisés avec peerScholar peuvent se terminer à cette étape et c'est souvent ce qui se produit. D'après les recherches effectuées par notre laboratoire (Paré et Joordens 2008) et également d'autres études (Cho, Schunn et Wilson 2006), une note finale établie à l'aide d'au moins cinq ou six notes attribuées par des pairs est aussi fiable que si elle avait été donnée par une aide-enseignante ou un aide-enseignant. Voilà qui confirme qu'il est possible d'exercer tous les résultats d'apprentissage signalés jusqu'ici sans avoir besoin de ressources pour la notation, ce qui est véritablement impressionnant<sup>3</sup>. Ce « processus partiel de peerScholar » et le schéma englobant les résultats d'apprentissage sont illustrés dans la figure 1.

---

<sup>3</sup> Même s'il a été démontré que l'utilisation de la moyenne des notes attribuées par les pairs était une méthode valable pour établir une note finale et malgré la richesse de l'apprentissage qui découle de ce processus, dans certains établissements avec des aides-enseignantes et aides-enseignants syndiqués (p. ex. l'Université de Toronto), on a réussi à faire reconnaître que c'était contraire à la législation sur le travail, même si les heures de travail n'étaient pas réduites. Dans les cas semblables, il faut employer le « processus complet de peerScholar », qui exige des ressources pour la notation, mais pas plus que pour une dissertation ordinaire, mais il en résulte alors une expérience d'apprentissage bien plus enrichissante.

**Figure 1 : Étapes du processus partiel de peerScholar pour les travaux scolaires d'après les résultats d'apprentissage de base courants**



Par contre, lorsqu'on a accès à un expert ou une experte pour l'évaluation finale, il est possible d'aller encore plus loin de manière à exercer encore plus les objectifs d'apprentissage critiques tout en permettant d'évaluer un plus grand nombre sur le plan quantitatif de la façon qui sera décrite plus loin. Si on applique le « processus complet de peerScholar » (figure 2), l'étudiante ou l'étudiant doit présenter une version révisée de son travail qui tient compte des commentaires faits par les pairs. La grande différence est à la troisième étape de l'organigramme. Après avoir lu et évalué les commentaires, l'étudiante ou l'étudiant doit faire deux choses : premièrement, réviser son travail en faisant de son mieux pour l'améliorer en fonction des commentaires reçus et, deuxièmement, rédiger un court texte expliquant pourquoi des changements ont été apportés et d'autres ne l'ont pas été à la lumière des commentaires précis qui ont été faits. Il faut dire aux étudiantes et étudiants qu'il est tout à fait correct de ne pas faire tous les changements recommandés, à condition de pouvoir justifier leur choix. On peut donc dire que cette étape et le processus de peerScholar en général illustrent le type de pratique de réflexion dont D.A. Shon et C. Argyris (1974 et 1978) font la promotion puisqu'on demande aux étudiantes et étudiants de mettre leur apprentissage en pratique de façon répétitive. C'est ce qu'on appelle parfois une « évaluation formative » et, encore une fois, de multiples recherches montrent que c'est une démarche qui contribue grandement à l'apprentissage (p. ex. Sadler 1989).

Ce travail de révision et de réflexion fait appel aux habiletés d'auto-réflexion et de communication réceptive. C'est déjà quelque chose de noter la valeur des suggestions qui nous sont faites, mais c'est autre chose que d'être capable de passer ces commentaires en revue et de s'en servir pour prolonger une idée ou un exposé de manière à l'améliorer. Par conséquent, dans le processus complet, l'étudiante ou l'étudiant doit, à la dernière étape, améliorer son propre travail en fonction de son analyse des commentaires formulés par les pairs.

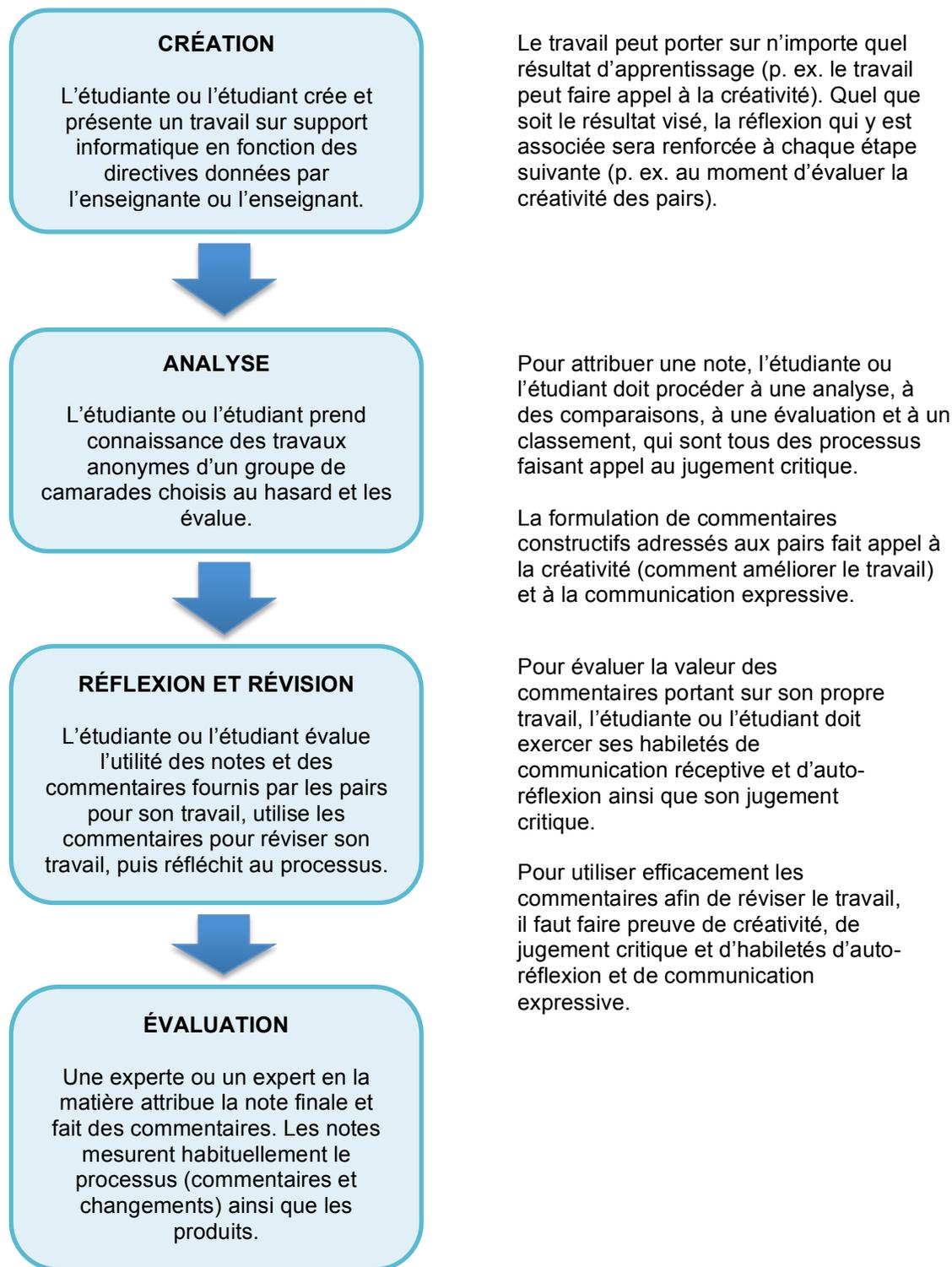
## Quantification des résultats d'apprentissage

Pour reprendre l'analogie de départ, l'objectif ultime du présent document est de prendre peerScholar comme exemple de la manière dont les habiletés cognitives associées aux résultats d'apprentissage essentiels peuvent être exercées et évaluées simultanément, tout comme les flexions des avant-bras permettent d'accroître et d'évaluer la force des biceps. La section précédente portait avant tout sur le volet « exercice » et espérons qu'elle a pu montrer de façon claire et convaincante à quel point les travaux effectués avec peerScholar permettent de s'exercer à atteindre pratiquement tous les résultats d'apprentissage essentiels de façon efficace et interreliée. Il faut dire que les diverses tâches ne servent pas à exercer seulement un résultat ou un autre mais plusieurs à la fois, un peu comme si on faisait des flexions des avant-bras tout en faisant du jogging! Même s'il est important de s'exercer, l'évaluation est tout aussi importante. Par conséquent, nous examinons maintenant la façon d'effectuer une évaluation quantitative des résultats d'apprentissage à l'aide de peerScholar.

Il faut commencer par faire la distinction entre les résultats qui sont mesurés automatiquement par le système et ceux qui sont quantifiés par un expert ou une experte en la matière (soit une aide-enseignante ou un aide-enseignant ou encore l'enseignante ou l'enseignant). Les résultats de la première catégorie peuvent être mesurés, que le processus complet de peerScholar soit employé ou non, mais ceux de la deuxième catégorie ne peuvent l'être que dans le processus complet. Nous commencerons par voir les résultats mesurés par un expert ou une experte en la matière pour ensuite passer aux possibilités qu'offre le système.

Dans le processus complet de peerScholar, la dernière étape comprend l'attribution d'une note par l'expert ou l'experte en la matière, et cette note est une mesure quantitative associée aux objectifs d'apprentissage de base. Un bon système de notation permet nécessairement de noter à la fois le produit de l'exercice et le processus sous-jacent. Quatre éléments sont généralement notés : 1) la qualité du travail initial, 2) la qualité du travail final, 3) l'utilité des commentaires faits aux pairs et 4) le bien-fondé des changements apportés.

**Figure 2 : Étapes du processus complet de peerScholar pour les travaux scolaires d'après les résultats d'apprentissage de base courants**



Vous vous rappelez que nous avons déjà dit que, si l'enseignante ou l'enseignant souhaite qu'un travail vise un résultat d'apprentissage en particulier, comme la responsabilité sociale, il ou elle peut demander un travail portant sur cet objectif. Dans ce cas, les deux premières notes attribuées par l'expert ou l'experte en la matière servent essentiellement à mesurer le niveau de l'étudiante ou de l'étudiant et ses progrès à court terme par rapport à ce résultat d'apprentissage, du moins d'après ce qu'on peut constater à partir du moyen employé pour faire le travail (p. ex. écriture, œuvre d'art, vidéo, etc.). Les deux autres « notes » expriment la capacité de l'étudiante ou de l'étudiant d'exercer un jugement critique, de faire preuve de créativité et de communiquer ses idées sur des améliorations possibles à ses pairs (utilité des commentaires faits aux pairs) ainsi que sa capacité de comprendre la communication des pairs, de porter un jugement critique et de procéder à une auto-réflexion pour ensuite faire une bonne révision de son travail (bien-fondé des changements apportés). Toutes ces relations entre les notes pour les divers volets et les objectifs d'apprentissage de base sont résumés dans la première partie du tableau 1.

**Tableau 1 : Critères d'évaluation quantitatifs à partir de peerScholar et relation avec les objectifs d'apprentissage de base**

<b>Critère de mesure</b>	<b>Objectif d'apprentissage connexe</b>
<b>Processus complet de peerScholar</b>	
<b>Qualité du travail initial</b>	Communication (sous la forme demandée) en plus de tout résultat d'apprentissage visé par le travail
<b>Qualité du travail final</b>	Communication (sous la forme demandée) en plus de tout résultat d'apprentissage visé par le travail
<b>Utilité des commentaires faits aux pairs</b>	Jugement critique (définir le principal problème), créativité (trouver quelles améliorations apporter) et communication (efficace et utile)
<b>Bien-fondé des changements apportés</b>	Communication réceptive (comprendre les commentaires des pairs), auto-réflexion et jugement critique (évaluer l'utilité des commentaires) et communication (adapter le travail en conséquence et justifier ses choix dans un document explicatif)
<b>Processus partiel de peerScholar</b>	
<b>Utilité des commentaires faits aux pairs (note moyenne pour l'utilité des commentaires)</b>	Jugement critique (définir le principal problème), créativité (trouver quelles améliorations apporter) et communication (efficace et utile)
<b>Qualité du travail (moyenne des notes attribuées par les pairs)</b>	Communication (sous la forme demandée) en plus de tout résultat d'apprentissage visé par le travail
<b>Dans les deux processus</b>	
<b>Écart par rapport aux évaluations des pairs (écart absolu moyen entre les notes)</b>	Jugement critique
<b>Écart de l'autoévaluation (note de l'autoévaluation comparativement à la note réelle)</b>	Auto-réflexion

Il est néanmoins aussi possible d'obtenir une évaluation quantitative des objectifs d'apprentissage de base même lorsqu'il n'y a pas d'expert ou d'experte en la matière qui donne des notes. Une fois le processus terminé, après l'évaluation des commentaires, l'étudiante ou l'étudiant ne modifie pas son travail et ne procède pas à une réflexion à son sujet, mais la moyenne des notes attribuées par les pairs devient sa note pour le travail. Encore une fois, la moyenne des notes des pairs traduit à la fois les habiletés de communication de la personne et son niveau de compétence par rapport à l'objectif d'apprentissage à la base du travail. Dans le processus partiel de peerScholar, c'est le seul moyen de mesurer les deux compétences à la fois, tandis que dans le processus complet, c'est une donnée qui aide l'expert ou l'experte en la matière à se faire une opinion sur la qualité du travail, mais il ou elle se base aussi sur sa propre évaluation du travail en question. C'est donc dire que le processus complet procure une évaluation plus riche mais cible les mêmes résultats de base.

De même, comme l'indique la partie du milieu du tableau 1, le processus partiel de peerScholar peut procurer une évaluation moins riche mais néanmoins valide de la qualité des commentaires faits aux pairs. N'oublions pas que chaque commentaire transmis reçoit une cote basée sur « l'utilité » de la part des pairs et que ces cotes peuvent être converties en chiffres pour obtenir une moyenne qui servira d'évaluation quantitative de la capacité des étudiantes et étudiants de porter un jugement critique pour déceler le principal problème, de faire preuve de créativité pour indiquer comment les pairs pourraient régler le problème, puis de communiquer leurs idées de façon claire, efficace et utile. Encore une fois, dans le processus complet de peerScholar, l'expert ou l'experte en la matière pourrait se baser sur cette évaluation en plus de ses propres impressions subjectives sur les commentaires, mais il ou elle évaluerait les mêmes habiletés cognitives.

Par conséquent, les deux processus de peerScholar peuvent procurer une évaluation quantitative des objectifs d'apprentissage de base en plus de permettre de s'exercer à atteindre les mêmes objectifs. Comme on l'a expliqué, il s'agit d'une évaluation composite, dans le sens qu'elle ne mesure pas simplement des objectifs d'apprentissage isolés mais plutôt différentes combinaisons d'habiletés cognitives. C'est en quelque sorte probablement inévitable. Dans le monde du travail, la réalisation des tâches exige habituellement l'utilisation dynamique et synergique d'une combinaison d'habiletés. Étant donné que les évaluations portent sur la réalisation des tâches, il n'y a rien d'étonnant à ce qu'elles mesurent la combinaison des habiletés. Cela dit, il serait fantastique de pouvoir obtenir des évaluations distinctes.

En réalité, il y a deux autres critères de mesure appliqués automatiquement dans peerScholar qui peuvent procurer une évaluation distincte du jugement critique et de l'auto-réflexion, que ce soit dans le processus partiel ou le processus complet. Il serait bien entendu naïf de croire que n'importe quel critère permette de bien mesurer toutes les facettes de concepts aussi complexes que le jugement critique et l'auto-réflexion, et il serait donc plus exact de dire qu'on évalue certaines facettes de ces résultats. Quoi qu'il en soit, il est pour le moins possible d'employer des travaux qui permettent à la fois de s'exercer à obtenir les résultats d'apprentissage et d'évaluer ces résultats. Ces critères de mesure figurent dans le bas du tableau 1.

Examinons d'abord le jugement critique. Le terme est utilisé de différentes façons dans les recherches, mais la plupart des gens conviennent que le concept renvoie, du moins en partie, à la capacité de procéder à une analyse, à des comparaisons et à une évaluation dans l'intention de vérifier la qualité d'une argumentation. Bien entendu, c'est exactement ce qu'on demande aux étudiantes et étudiants de faire durant l'étape d'évaluation d'un travail réalisé à l'aide de peerScholar, et la note qu'ils attribuent traduit leur opinion sur la qualité d'un travail donné. Il s'avère qu'en faisant certaines hypothèses raisonnables, il est possible de

quantifier l'exactitude des notes attribuées par une étudiante ou un étudiant et, de cette manière, se faire une idée de sa capacité d'exercer un jugement critique et de vérifier la qualité du travail évalué.

L'évaluation du jugement critique se fait de la façon qui suit. Une étudiante ou un étudiant donné évalue généralement six pairs, qui sont chacun évalués par cinq autres étudiantes et étudiants. Si on suppose que la moyenne des notes attribuées par les cinq autres personnes représente une bonne évaluation de la qualité réelle du travail (Paré et Joordens 2008) et qu'on calcule l'écart entre la note donnée par une étudiante ou un étudiant en particulier et la moyenne des notes données par les autres, qu'on prend l'écart absolu pour ensuite additionner tous les écarts absolus et les diviser par le nombre d'écarts, on obtient ce qui s'appelle un écart absolu moyen. Voici comme s'illustre cette formule.

$$\text{Jugement critique} = \frac{\sum_1^p |s - \bar{x}_g|}{p} \quad [1]$$

$p$  = le nombre de travaux notés par des pairs

$s$  = la note qu'une étudiante ou un étudiant donne à un travail précis

$\bar{x}_g$  = la moyenne des notes données par les autres étudiantes et étudiants pour le même travail

À signaler que ce qu'on obtient est essentiellement un score d'erreur, dans le sens que plus la note attribuée par une étudiante ou un étudiant diffère de la moyenne des notes attribuées par les autres, plus le score est élevé. Autrement dit, une étudiante ou un étudiant doté d'un excellent jugement critique devrait normalement attribuer des notes très proches de la moyenne des notes des autres, et l'écart moyen devrait être faible dans l'ensemble. Si on préfère que ce soit un score élevé qui indique un bon jugement critique, il suffit d'inverser les chiffres.

Le calcul à effectuer pour l'auto-réflexion doit, au contraire, être basé principalement sur les scores d'autoévaluation qu'on peut demander à l'étudiante ou à l'étudiant de fournir à n'importe quelle étape du processus de peerScholar. Encore une fois, si on suppose que la moyenne des notes attribuées par les pairs pour le travail est le reflet fidèle de la qualité réelle du travail (Paré et Joordens 2008), on peut simplement calculer l'écart entre la note de l'autoévaluation et la moyenne des notes des pairs et tenir pour acquis que plus l'écart est petit, plus les habiletés d'auto-réflexion sont bonnes.

Dans ce cas aussi, le résultat exprime le degré d'erreur puisqu'un score élevé signifie que l'étudiante ou l'étudiant ne sait pas bien juger la qualité de son propre travail. À noter que, vu qu'on peut demander aux étudiantes et étudiants de faire une autoévaluation aux étapes de création, d'évaluation et de réflexion et de révision, il est possible de faire le suivi des progrès des habiletés d'auto-réflexion durant l'exécution d'un travail.

## Sommaire de la section

L'objectif de cette section était de définir de façon détaillée comment un outil conçu pour s'exercer par rapport aux résultats d'apprentissage pouvait aussi servir à évaluer ces mêmes résultats. Les évaluations obtenues ne sont pas toutes « pures », et nous n'en sommes donc pas encore à pouvoir obtenir une ventilation de chaque résultat d'apprentissage particulier, ce qui n'est peut-être pas un objectif réaliste puisque les habiletés cognitives à la base de ces résultats doivent souvent être combinées pour résoudre un problème. Il est

néanmoins certain qu'on peut arriver à faire comme dans l'analogie avec un programme d'entraînement, soit s'exercer et s'évaluer à la fois en effectuant des travaux.

## Des données probantes

Maintenant qu'on a établi qu'il était possible de s'exercer à atteindre les résultats d'apprentissage et de les évaluer du même coup, il serait bon, pour être le plus convaincant possible, de savoir que les évaluations dont on a déjà parlé montrent qu'il y a des progrès qui sont réalisés au fil du temps. Lorsque les étudiantes et étudiants évaluent leur travail à répétition, leurs habiletés d'auto-réflexion s'améliorent-elles? À force d'exercer leur jugement critique, deviennent-ils plus efficaces? Voilà des questions auxquelles on a tenté de répondre directement en faisant les deux expériences indiquées ci-dessous.

### Première expérience : auto-réflexion

Comme on l'a déjà dit, on peut demander aux étudiantes et étudiants qui font un travail à l'aide de peerScholar de faire une évaluation quantitative de la qualité de leur propre travail à la fin de n'importe laquelle des trois étapes. Cette autoévaluation peut être comparée à une « note finale réelle », qu'il s'agisse de la moyenne des notes attribuées par des pairs pour le travail ou de la note donnée par une experte ou un expert, selon qu'on a accès ou non à des experts pour la note finale. La question essentielle à laquelle l'expérience devait permettre de répondre était de savoir si l'exactitude des autoévaluations augmente, ou plutôt si l'écart entre les autoévaluations et la note finale réelle diminue, en fonction de l'exposition aux travaux des pairs et de leur évaluation.

### Personnes participantes

Ce sont les 1 292 étudiantes et étudiants du cours d'introduction à la psychologie de l'automne 2010 (A01) offert par l'Université de Toronto à Scarborough qui ont consenti à ce que les données relatives à leur travail soient utilisées pour l'étude.

### Travail et documents présentés

Le travail en question consistait à rédiger une argumentation pour ou contre la pratique actuelle qui est d'appliquer aux animaux des protocoles de recherche qui sont jugés contraires à l'éthique pour les êtres humains. Il s'agissait donc de créer, puis d'évaluer et de réviser de courtes argumentations (d'environ une à deux pages). Tous les documents ont été présentés dans peerScholar, comme on l'a déjà mentionné dans le présent rapport.

### Démarche

On a demandé à l'ensemble des étudiantes et des étudiants répartis dans deux groupes de faire un travail à l'aide de peerScholar, mais les particularités du travail n'étaient pas les mêmes selon le groupe. Pour cette étude, le processus partiel de peerScholar a été employé dans les deux groupes (c.-à-d. que le travail n'avait pas à être révisé), et la seule différence était la fréquence des autoévaluations. On a demandé à la moitié des étudiantes et des étudiants (soit 646) de faire une autoévaluation après avoir remis leur première version

durant l'étape de création et après avoir évalué le travail des pairs à l'étape d'évaluation, tandis que les autres n'ont eu à faire qu'une seule autoévaluation durant l'étape d'évaluation. Les résultats obtenus pour le deuxième groupe ne sont pas pertinents pour les besoins du présent rapport, mais ils étaient néanmoins inclus dans le projet de recherche. Puisque le but premier est ici d'évaluer les progrès des autoévaluations en fonction des évaluations des pairs, nous nous en tiendrons au premier groupe.

Les personnes faisant partie du groupe ayant subi la « double évaluation » ont commencé par composer leur première version et, après l'avoir remise, elles devaient simplement se donner une note sur 10 correspondant à leur évaluation de la qualité du travail. Par la suite, à l'étape d'évaluation, elles ont d'abord évalué, examiné et commenté le travail de six pairs, puis elles ont dû réévaluer leur propre travail après avoir reçu la directive suivante : « Vous vous rappelez sans doute la note que vous avez donnée à votre travail à l'étape précédente, mais essayez d'évaluer maintenant votre travail comme si vous le voyiez pour la première fois. Compte tenu de ce que vous avez appris en notant les travaux de vos pairs, examinez votre travail initial encore une fois et indiquez la note qu'il mérite selon vous. »

### **Résultats et analyse**

Ce qu'il faut déterminer, c'est si le fait d'avoir été exposé aux travaux des pairs et de les avoir analysés a permis de réduire l'écart entre l'autoévaluation et la note réellement méritée. Autrement dit, en comparant la note estimative avec la moyenne des notes attribuées par les pairs pour le travail (Paré et Joordens 2008) et en calculant l'écart, qui est essentiellement le reflet d'un manque de connaissance de soi, on peut déterminer si les étudiantes et étudiants apprennent à mieux se connaître en voyant et en évaluant les travaux des pairs.

Il est intéressant de constater que les étudiantes et étudiants ont surestimé leur note tant avant qu'après l'évaluation par les pairs, c'est-à-dire que la moyenne des autoévaluations était de 8,14 à la fin de l'étape de création et de 7,74 à la fin de l'étape d'évaluation, ce qui est dans les deux cas supérieur à la moyenne des notes réelles, à 6,64. La moyenne à chaque étape était statistiquement significative comparativement à la moyenne des notes attribuées par les pairs d'après un test  $t$ , qui était approprié étant donné que les scores à la base de chaque moyenne étaient normalement distribués et que, dans chaque cas, ils étaient jugés, avec un degré de fiabilité élevé, supérieurs à la moyenne des notes attribuées par les pairs,  $t(644) > 15,23$ ,  $p < ,0001$ . Ce n'est là qu'un autre exemple du phénomène bien connu en psychologie sociale qu'on appelle « biais de complaisance », soit la tendance à croire que notre rendement est meilleur qu'il ne l'est en réalité afin de garder une bonne estime de soi (Miller et Ross 1975).

Ce qui est particulièrement pertinent pour le présent rapport, c'est que les écarts par rapport à l'autoévaluation avant et après l'évaluation par les pairs (soit 1,5 et 1,1 respectivement) ont aussi été comparés l'un à l'autre et que l'écart obtenu était aussi statistiquement fiable,  $t(644) > 7,80$ ,  $p < ,0001$ . Cette constatation montre qu'il suffit d'une seule évaluation de six pairs pour améliorer la capacité des étudiantes et étudiants d'évaluer la qualité de leur propre travail, c'est-à-dire leur habileté d'auto-réflexion.

### **Deuxième expérience : jugement critique**

Le concept du jugement critique comporte de multiples facettes, et il est un peu audacieux de croire qu'on puisse mesurer le jugement critique d'une étudiante ou d'un étudiant au moyen d'une évaluation quantitative relativement simple. Pourtant, quand on songe au type de processus cognitifs qu'il faut appliquer pour

évaluer la qualité de travaux, surtout lorsqu'il s'agit de travaux de pairs, les termes qui nous viennent à l'esprit sont ceux qui sont typiquement associés au concept du jugement critique, puisqu'on demande aux étudiantes et étudiants de faire une évaluation, des comparaisons, une analyse, un classement, etc. Par conséquent, même si cela peut sembler audacieux, nous estimons que le degré d'exactitude avec lequel une étudiante ou un étudiant est capable de noter une série de travaux différents donne un très bon aperçu, quoique simplifié, de son jugement critique. Si c'est vraiment le cas et si peerScholar aide réellement les étudiantes et étudiants à développer leur jugement critique, on peut s'attendre à ce que l'exactitude des notes qu'ils attribuent s'améliore aussi à mesure qu'ils prennent de l'expérience. Voilà la question à laquelle l'expérience a tenté de répondre.

### *Personnes participantes*

Les 1 279 étudiantes et étudiants du cours d'introduction à la psychologie à l'automne 2011 (A01) et à l'hiver 2012 (A02) offert par l'Université de Toronto à Scarborough ont consenti à ce que les données relatives à leur travail soient utilisées pour les besoins de l'étude. De ce nombre, 737 ont fait un travail à l'aide de peerScholar aux périodes 1, 2 et 3 (groupe d'exercices répétés), tandis que les 542 autres ont fait un seul travail à l'aide de peerScholar à la période 3 (groupe de contrôle). Il est possible de suivre les deux premières parties du cours d'introduction à la psychologie séparément et dans n'importe quel ordre. Ceux et celles ayant suivi la partie A01 puis la partie A02 ont fait leur troisième travail à l'aide de peerScholar à la période 3, tandis que ceux et celles ayant suivi la partie A02 mais pas la partie A01 ont fait leur seul travail à l'aide de peerScholar à cette période.

### *Travail et documents présentés*

Le travail en tant que tel était différent aux trois périodes, mais chaque fois, les étudiantes et les étudiants ont dû rédiger une brève argumentation convaincante pour ou contre un sujet ou une pratique. On a veillé à ce que le niveau de difficulté demeure le même. Tous les documents ont été présentés dans peerScholar, comme on l'a déjà mentionné dans le présent rapport, et chacun a été évalué par six pairs.

### *Démarche*

Au moment de l'étude, les étudiantes et étudiants faisant la première partie du cours d'introduction à la psychologie ont effectué deux travaux à l'aide de peerScholar, tandis que ceux et celles suivant la deuxième partie ont fait un seul travail. On a appliqué le processus partiel de peerScholar déjà décrit et, d'après les évaluations fournies par les étudiantes et étudiants, on a pu calculer l'écart moyen absolu pour le jugement critique conformément à l'équation [1]. C'est donc dire que le groupe d'exercices répétés a fourni trois scores, soit un à chaque niveau de la variable temporelle, alors que le groupe de contrôle n'a fourni qu'un seul score associé à la période 3.

### *Résultats et analyse*

L'expérience avait pour but d'évaluer l'influence de peerScholar sur les scores relatifs au jugement critique en se basant sur deux questions. Premièrement, les scores diminuent-ils avec la pratique, comme on pourrait le croire d'après l'analogie avec un programme d'entraînement? Il est possible de répondre à cette question simplement en examinant les scores du groupe d'exercices répétés aux trois périodes. Deuxièmement, si on

observe effectivement une diminution, peut-on être relativement certain que celle-ci est attribuable à la pratique et non pas simplement au fait de se trouver dans un milieu universitaire pendant un trimestre? Une comparaison entre le groupe d'exercices répétés et le groupe de contrôle permet d'obtenir une réponse relativement certaine à cette question, vu que les deux groupes ont été dans un milieu universitaire pendant au moins un trimestre mais que seul le groupe d'exercices répétés a déjà utilisé peerScholar.

Pour ce qui est de la première question, une analyse de la variance à un critère de classification a été effectuée avec la variation des scores de l'écart absolu moyen (calculés à l'aide de l'équation [1]) comme variable dépendante et la période où a été effectué le travail pour le groupe d'évaluation par les pairs (période 1, 2 ou 3) comme variable indépendante. Un écart significatif a été observé selon la période où s'est fait le travail,  $F(2, 1\ 472) = 24,7$ ,  $p < 0,001$ ,  $\eta^2 = 0,03$ . Reportez-vous au tableau 1, qui contient chaque facteur. Compte tenu de ce résultat significatif et d'après notre hypothèse, nous avons fait un examen plus poussé des données et nous avons constaté une tendance linéaire significative,  $F(2, 1\ 472) = 24,7$ ,  $p < 0,001$ ,  $\eta^2 = 0,06$ .

**Tableau 2 : Scores de l'écart absolu moyen pour le jugement critique (les scores plus bas sont le signe d'une plus grande concordance avec la note finale)**

	Période 1	Période 2	Période 3
Exercices répétés	1,08	0,98	0,95
Contrôle			1,04

Ces résultats montrent clairement que la participation à une évaluation par les pairs contribue à améliorer la capacité des étudiantes et des étudiants de mesurer la qualité. Non seulement la première participation à l'étape d'évaluation des travaux a permis de perfectionner les habiletés d'évaluation de la qualité, mais chaque participation par la suite a eu le même effet. Voilà qui confirme notre hypothèse selon laquelle une exposition répétée à des évaluations par les pairs contribue à améliorer la capacité des étudiantes et des étudiants de reconnaître la qualité et donc leur jugement critique.

Nous avons ensuite comparé les scores de l'écart absolu moyen du groupe de contrôle pour le premier travail (période 3) à ceux du groupe d'exercices répétés pour le premier travail (période 1) et pour le troisième travail (période 3). Entre la période 3 pour le groupe de contrôle et la période 1 pour le groupe d'exercices répétés, il n'y avait pas d'écart significatif entre les scores de l'écart absolu moyen,  $t(1\ 277) = -1,876$ ,  $p = 0,061$ , n.s., tandis que pour la période 3, l'écart entre le groupe de contrôle et le groupe d'exercices répétés était significatif, soit  $t(1\ 277) = 3,573$ ,  $p < 0,001$ .

Par conséquent, comme pour l'auto-réflexion, le jugement critique progresse avec l'utilisation de peerScholar, comme on peut s'y attendre si l'analogie avec le programme d'entraînement tient toujours, c'est-à-dire qu'à force de s'exercer à évaluer les travaux de pairs, les étudiantes et étudiants apprennent à mieux évaluer la qualité des travaux et, dans la mesure où l'évaluation de la qualité découlant des comparaisons, de l'examen, de l'analyse et du classement est le reflet du jugement critique, il semblerait que le jugement critique, ou du moins un indice qui y est associé, puisse à la fois être exercé et évalué en faisant un seul et même travail.

## Sommaire

Il est relativement facile de définir les résultats d'apprentissage, mais il est beaucoup plus difficile d'expliquer dans le détail comment s'exercer à les atteindre et les évaluer. D'aucuns ont suggéré d'utiliser des données globales subjectives capables de rendre compte de l'expérience des étudiantes et étudiants sur une longue durée, comme celles que permettraient de recueillir des sondages sur l'expérience des étudiantes et étudiants. À notre avis, il serait bien plus souhaitable de s'exercer à atteindre les résultats d'apprentissage et de les évaluer un travail à la fois et de partir de là pour faire des extrapolations.

Le but premier du présent rapport était de faire la preuve que c'était là un objectif réalisable. En concentrant notre étude sur un outil technologique en particulier, soit peerScholar, nous avons montré comment des données quantitatives d'évaluation relatives aux résultats d'apprentissage pouvaient être compilées pour un travail à la fois. Nous avons même présenté de récentes données de notre laboratoire, qui indiquent que les résultats d'évaluation sont conformes aux attentes, c'est-à-dire qu'on observe des améliorations avec la pratique cognitive. Voilà qui prouve, avec un degré élevé de certitude, qu'il est bel et bien possible de s'exercer et de procéder à une évaluation pour un travail à la fois.

Signalons que nous ne prétendons pas ici que tout le monde devrait se servir de peerScholar, quelle que soit notre opinion sur la question. Nous avons fait porter notre étude sur cet outil parce que c'est celui que nous connaissons le mieux, puisqu'il a été conçu par notre laboratoire. D'autres analyses semblables pourraient et devraient être faites dans d'autres contextes relatifs aux travaux scolaires, ce qui pourrait déboucher sur la création d'un vaste éventail d'outils dont on pourrait se servir tant pour s'exercer que pour évaluer divers objectifs d'apprentissage de différentes façons. Idéalement, on pourrait élaborer des exercices servant à prouver la maîtrise d'objectifs d'apprentissage donnés et en faire le suivi de façon officielle, notamment avec des outils comme Mozilla Badges ou des portfolios électroniques, et ces exercices pourraient finir par devenir plus pertinents que les notes traditionnelles. Mais pour en arriver là, il faut d'abord relier les travaux aux résultats d'apprentissage et, avec de la chance, le présent rapport a démontré de manière évidente que c'était chose possible. Il est temps de passer à l'action et de commencer le programme d'entraînement des habiletés.

## Bibliographie

- CHO, K., C.D. SCHUNN et R.W. WILSON. "Validity and Reliability of Scaffolded Peer Assessment of Writing from Instructor and Student Perspectives", *Journal of Educational Psychology*, vol. 98, 2006, p. 891-901.
- CRAIK, F.I.M., et R.S. LOCKHART. "Levels of Processing: A Framework for Memory Research", *Journal of Verbal Learning and Verbal Behavior*, vol. 11, 1972, p. 671-684.
- FOUNDATION FOR CRITICAL THINKING. 2013. [<http://www.criticalthinking.org/pages/critical-thinking-where-to-begin/796>].
- JOORDENS, S., S. DESA et D.E. PARÉ. "The Pedagogical Anatomy of Peer-Assessment: Dissecting a peerScholar Assignment", *The Journal on Systemics, Cybernetics and Informatics*, vol. 7, n° 5, 2009, p. 11-15.
- KING, A. "Structuring Peer Interaction to Promote High-Level Cognitive Processing", *Theory into Practice*, vol. 41, 2002, p. 33-39.
- MILLER, D.T., et M. ROSS. "Self-Serving Biases in the Attribution of Causality: Fact or Fiction?", *Psychological Bulletin*, vol. 82, n° 2, 1975, p. 213-225.
- MILNER, B.; Cité dans PASSOUANT, P. (éd.). *Physiologie de l'hippocampe*, Paris, Centre national de la recherche scientifique, 1962, p. 257-272.
- PARÉ, D.E., et S. JOORDENS. "Peering into Large Lectures: Examining Peer and Expert Mark Agreement Using peerScholar, an Online Peer-Assessment Tool", *Journal of Computer Assisted Learning*, vol. 24, n° 6, 2008, p. 526-540.
- SADLER, D.R. "Formative Assessment and the Design of Instructional Systems", *Instructional Science*, vol. 18, 1989, p. 119-144.
- SAMUELS, S.J., et R.F. FLOR. "The Importance of Automaticity for Developing Expertise in Reading", *Reading & Writing Quarterly*, vol. 13, n° 2, 1997, p. 107-121.
- SCHON, D.A., et C. ARGYLIS. *Theory in Practice: Increasing Professional Effectiveness*, San Francisco, Californie, Jossey-Bass, 1974.
- SCHON, D.A., et C. ARGYLIS. *Organizational Learning: A Theory of Action Perspective*, Reading, Massachusetts, Addison-Wesley, 1978.
- TULVING, E. "How Many Different Memory Systems Are There?", *American Psychologist*, vol. 40, 1985, p. 385-398.
- VENABLES, A., et R. SUMMIT. "Enhancing Scientific Essay Writing Using Peer Assessment", *Innovations in Education and Teaching International*, vol. 40, 2003, p. 281-290.
- VYGOTSKY, L.S. *Mind in Society: The Development of Higher Psychological Processes*, Cambridge, Massachusetts, Harvard University Press, 1978.

...the ...

...the ...