

Journées du CEEIADE
22 et 23 mars 2023

Comment évaluer les compétences en stages? Et le raisonnement clinique?

André-Sébastien Aubin, Ph. D.
Section Méthodologie d'évaluation en éducation



UQÀM
Université du Québec à Montréal

1

Présentateur

UQÀM



Répertoire des professeurs



André-Sébastien Aubin

Professeur


 **Unité** : Département d'éducation et pédagogie

 **Courriel** : aubin.andre-sebastien@uqam.ca







 **Téléphone** : (514) 987-3000 poste 6405

 **Local** : N-6840

 **Langues** : Français, Anglais

 Ce professeur désire s'entretenir avec les médias

Domaines d'expertise

- Auxiliaires d'enseignement 
- Éducation 
- Évaluation des compétences 
- Modèles de mesure en éducation 
- Pédagogie en milieu universitaire 
- Validation en contexte de classe 

Journées du CEEIADE 22 et 23 mars 2023

2

2

Plan de la présentation

- Une introduction
 - Un petit glossaire de l'évaluation des apprentissages
 - Les trois piliers de l'**approche par compétences** (ApC)
 - Le fardeau de la qualité de l'**évaluation des apprentissages** (EdA)
- Évaluer dans le cadre de stages
- Évaluer le raisonnement clinique
- S'outiller pour réaliser les observations dans le contexte de stage
- La conclusion

Journées du CEELADE 22 et 23 mars 2023

3

3

Parler d'évaluation des apprentissages!



4

4

Un petit glossaire de l'EdA

L'évaluation des apprentissages (EdA) : L'évaluation des apprentissages est la démarche qui commence par l'identification des cibles d'apprentissages et se termine par la prise de décisions. Scallon (2004) proposait de reprendre le terme anglophone « Assessment » qui reprend cette idée de démarche ou de processus.

Dans la présentation, nous n'utiliserons jamais le mot évaluation comme un synonyme de tâche, mais bien selon le sens de processus menant à la prise de décision.

Cela exclut des phrases comme « Aujourd'hui les étudiantes et les étudiants font l'évaluation. » « C'est la semaine de l'évaluation » « C'est une évaluation formative. »

Un petit glossaire de l'EdA

Les décisions : Les décisions sont les jugements obtenus par le processus d'évaluation des apprentissages. Ces décisions peuvent porter sur le niveau de compétence de l'étudiante ou étudiant, sur les activités d'enseignement et d'apprentissage ou sur la sélection de certaines étudiantes ou certains étudiants.

Une tâche: Tout ce que l'on demande de faire aux étudiantes ou aux étudiants en lien avec l'EdA. Cela peut inclure les exposés oraux, les projets, les essais, les examens.

Les performances: C'est le terme générique décrivant ce que les étudiantes et les étudiants font en réponse à une tâche.

Les trois concepts fondamentaux de l'EdA dans l'ApC

L'ApC est l'adaptation dans la francophonie de plusieurs courants anglo-saxons.

L'EdA dans l'ApC repose sur trois piliers.



La mobilisation des connaissances, et non pas seulement leur mémorisation.



La trajectoire de développement.



La priorité de la fonction d'aide à l'apprentissage de l'évaluation.

Image: DALL-E (2023)

8

8



La mobilisation des compétences

Si l'on veut observer la capacité des étudiantes et des étudiants à mobiliser de façon efficace des ressources, il faut utiliser des tâches servant à l'évaluation des apprentissages qui leur demandent de le faire.

Ces tâches sont appelées des **tâches complexes**.

Les tâches complexes: Les tâches complexes sont des tâches qui demandent aux étudiantes ou aux étudiants de déterminer par eux-mêmes (avec ou sans aide) ce qu'ils doivent faire durant la tâche.

Les SAÉs: Les situations d'apprentissages et d'évaluation regroupent plusieurs tâches plus ou moins complexes. Les SAÉs qui permettent d'évaluer le niveau des étudiantes et des étudiants sur une même compétence font partie de la même famille de situation.

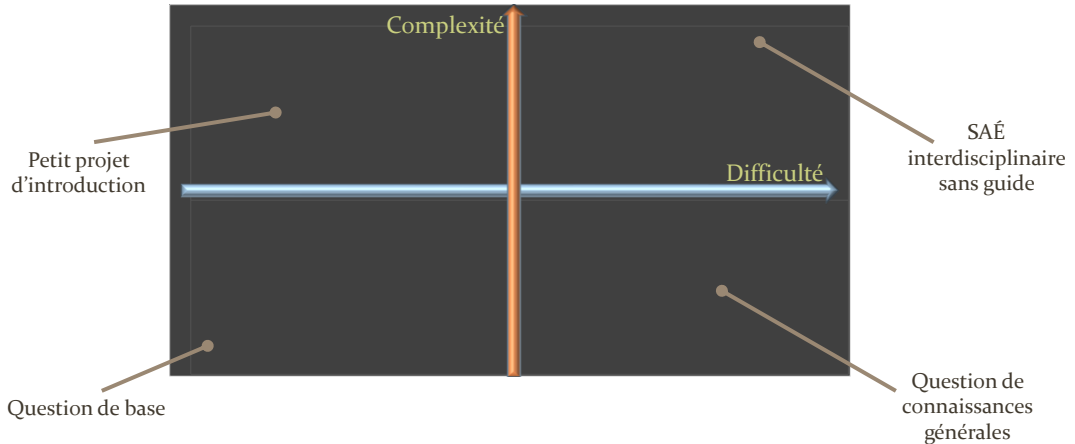


9



Les tâches complexes

Les tâches complexes demandent aux étudiantes et aux étudiants de résoudre un problème dont la solution n'est pas immédiatement apparente.



10



Les tâches complexes

Les principaux facteurs de complexité d'une tâche

- Le nombre « d'univers » touchés par la tâche.
- La nouveauté de la tâche.
- Le niveau des processus cognitifs nécessaires pour y répondre.

La complexité est une caractéristique d'une tâche utilisée à un moment précis.

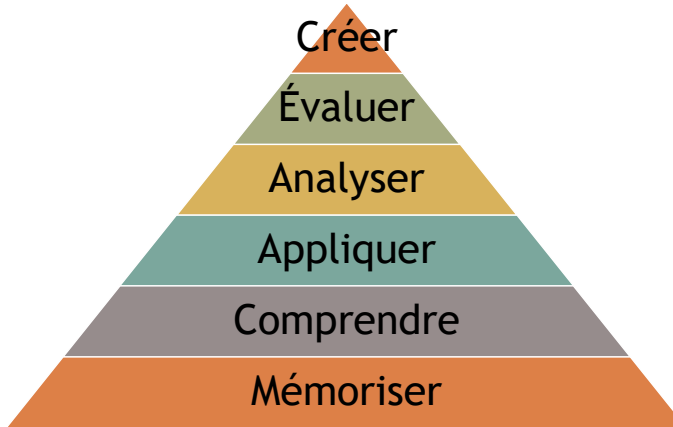
- Une tâche complexe au premier cycle peut ne plus l'être au second.

11



Les tâches complexes

Bloom propose une façon de classer les processus cognitifs



Version révisée



12



Les tâches complexes



Tarik Glenn - Offensive Tackle	
#78	6-5, 332 California 10 years in the NFL D1-97 DOB: 05/25/1976



Terrence Wilkins - Wide Receiver	
#10	5-10, 180 Virginia 6 years in the NFL FA-06 DOB: 07/29/1975



13



Les tâches complexes



Assessment is authentic when we directly examine student performance on worthy intellectual tasks. Traditional assessment, by contrast, relies on indirect or proxy 'items'--efficient, simplistic substitutes from which we think valid inferences can be made about the student's performance at those valued challenges.

Do we want to evaluate student problem-posing and problem-solving in mathematics? Experimental research in science? speaking, listening, and facilitating a discussion? doing document-based historical inquiry? thoroughly revising a piece of imaginative writing until it "works" for the reader? Then let our assessment be built out of such exemplary intellectual challenges.

– Wiggins, 1990



14



La pyramide de Scallon (2004)



16



La trajectoire de développement.

Si l'on veut observer la capacité des étudiantes et des étudiants à mobiliser de façon efficace des ressources, il faut utiliser des tâches servant à l'évaluation des apprentissages qui leur demandent de le faire.

Ces tâches sont appelées des **tâches complexes**.

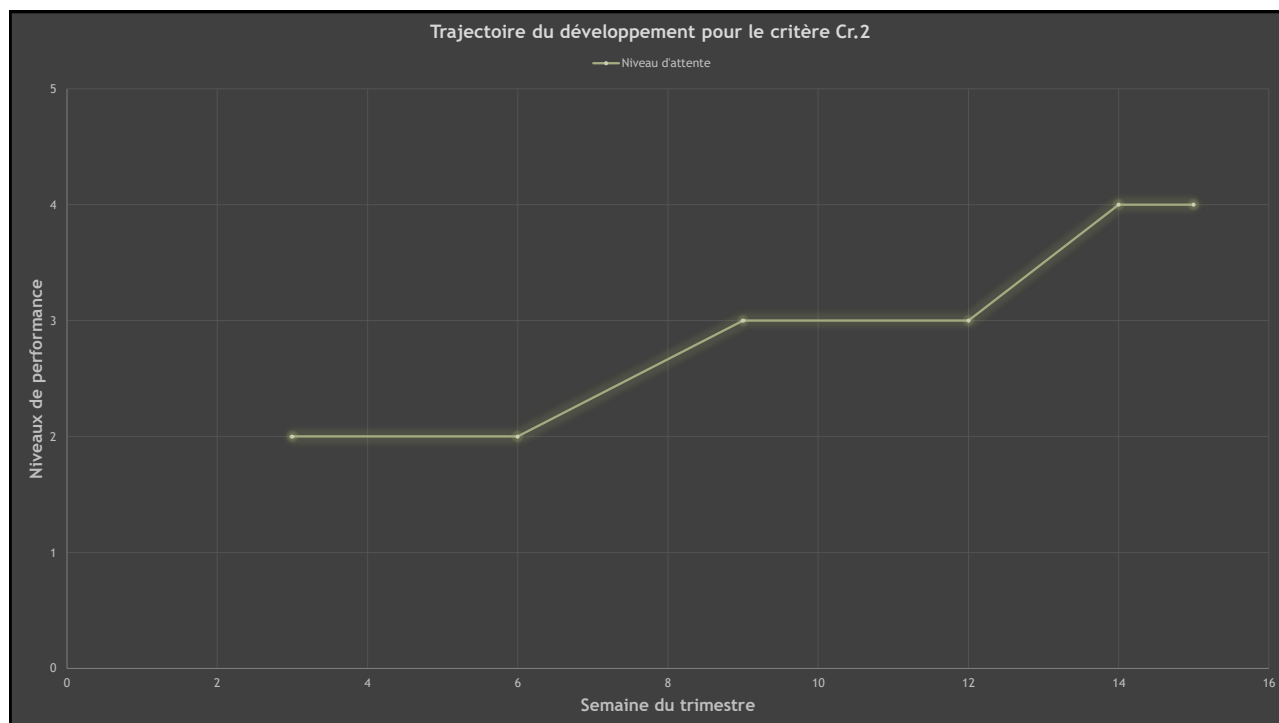
Une tâche: Tout ce que l'on demande de faire aux étudiantes ou aux étudiants en lien avec l'EdA. Cela peut inclure les exposés oraux, les projets, les essais, les examens.

Les tâches complexes: Les tâches complexes sont des tâches qui demandent aux étudiantes ou aux étudiants de déterminer par eux-mêmes (avec ou sans aide) ce qu'ils doivent faire durant la tâche.

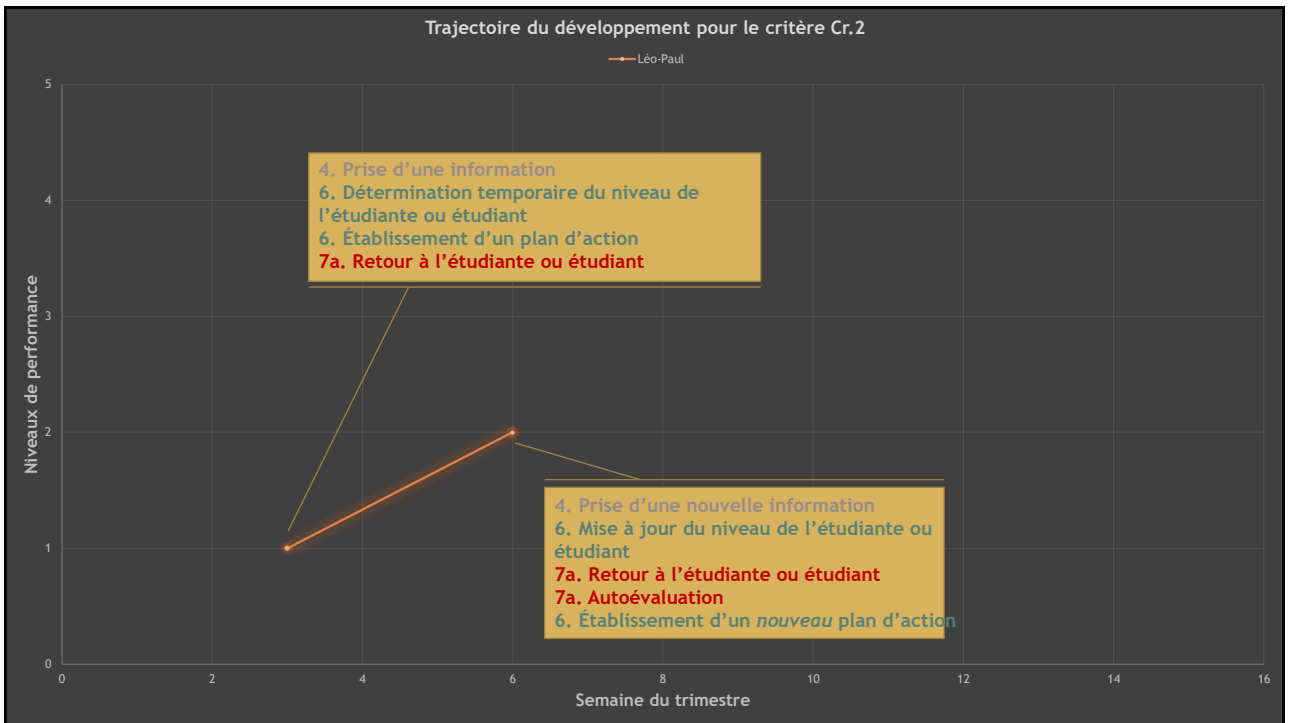
Les SAÉs: Les situations d'apprentissages et d'évaluation regroupent plusieurs tâches plus ou moins complexes. Les SAÉs qui permettent d'évaluer le niveau des étudiantes et des étudiants sur une même compétence font partie de la même famille de situation.



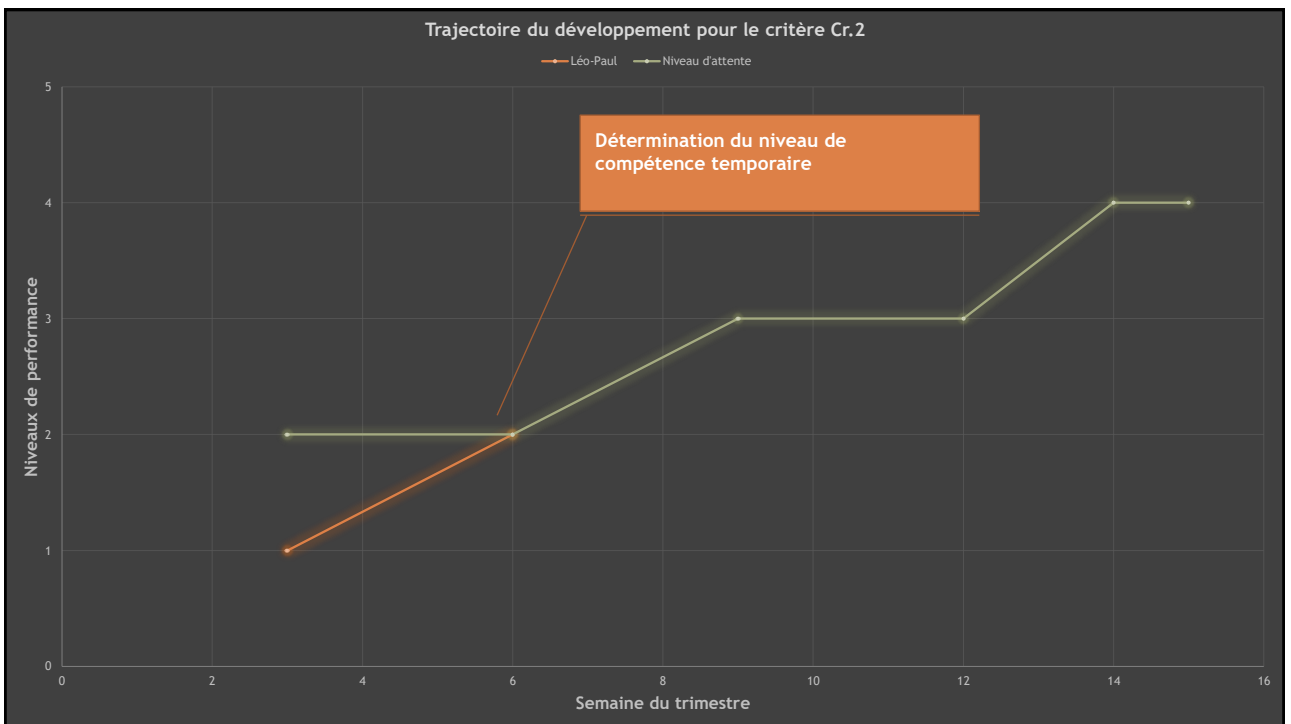
17



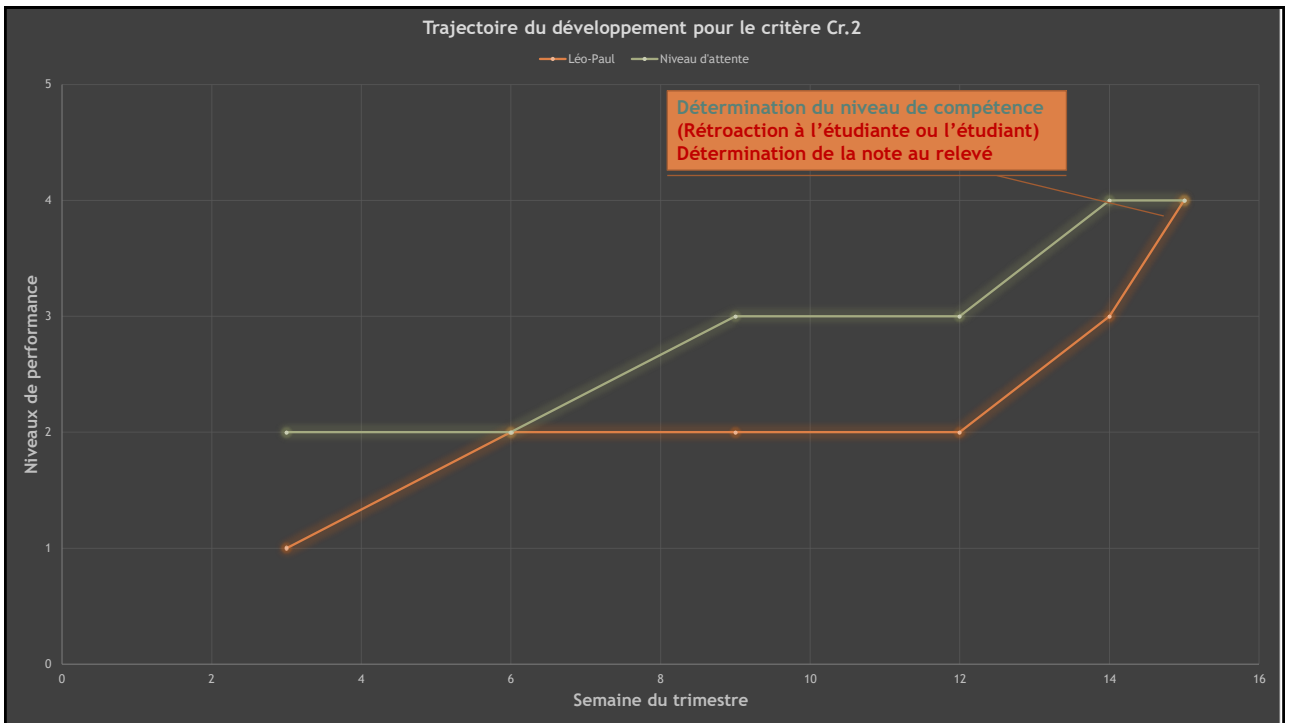
18



19



20



21



Des observations à la décision

La décision la plus importante est de déterminer le niveau de compétence de chaque étudiante et de chaque étudiant. Dans une approche analytique et critériée, on doit établir le niveau de compétence à chacune des tâches en compilant les observations des différents critères.

Pour cela, on doit déterminer une « politique » de jugement, c'est-à-dire la façon dont nous passons des observations à la décision.

22



Des observations à la décision

La politique de jugement décrit de quelle façon l'enseignant passera des informations qu'il a recueillies à un jugement sur le niveau de l'étudiante ou de l'étudiant.

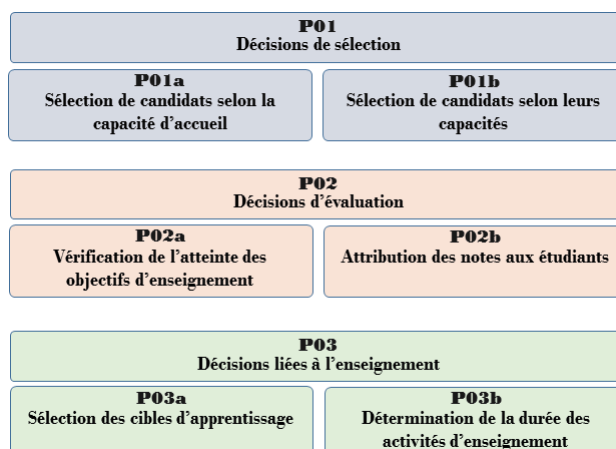
Cette politique peut varier selon les compétences et leurs importances relatives dans le programme. Les courants les plus fréquents en évaluation des compétences sont:

- Niveau maximal observé (risque de faux positifs)
- Niveau minimal observé (risque de faux négatifs)
- Moyenne pondérée par critère des niveaux observés (tient peu compte de la progression)



La fonction d'aide à l'apprentissage de l'évaluation

L'évaluation des apprentissages permet de prendre des décisions à partir des données recueillies. Popham (2000) propose une façon d'organiser les types de décisions pouvant être prises.





La fonction d'aide à l'apprentissage de l'évaluation

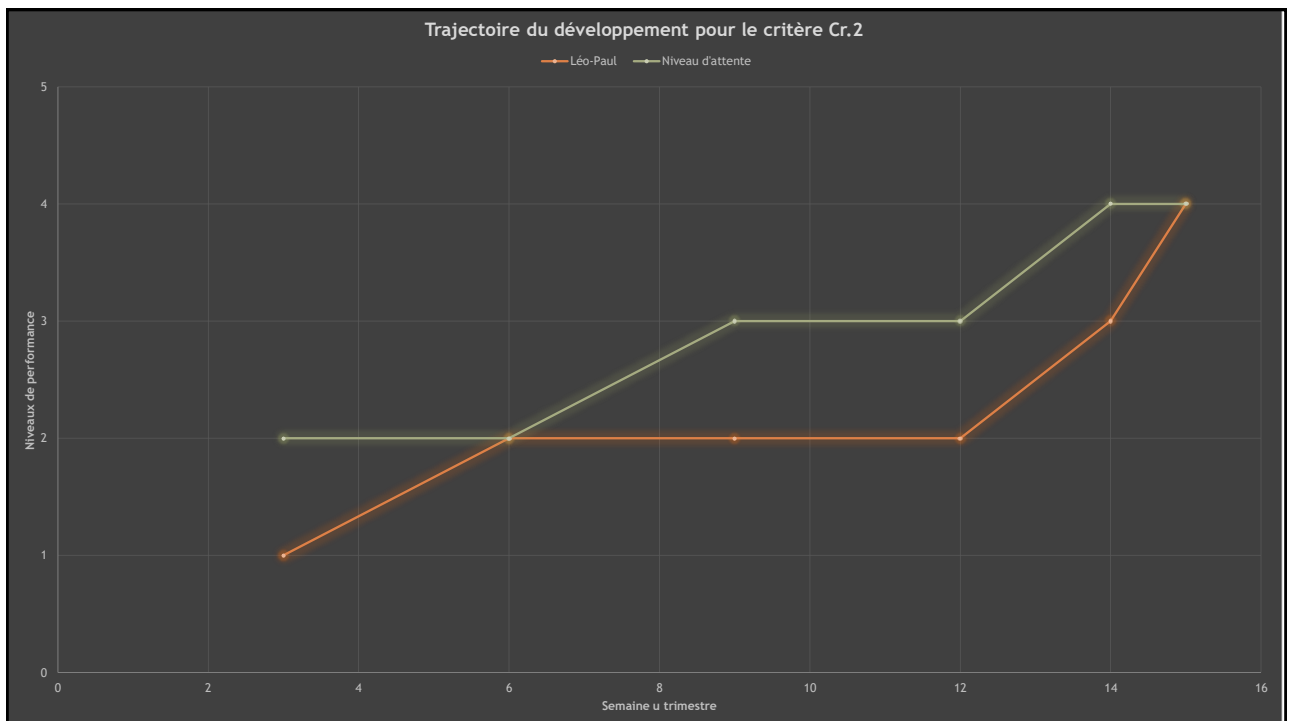
Selon le type de décisions, la fonction mise de l'avant n'est pas la même. Cependant, il arrive souvent que plusieurs fonctions soient recherchées simultanément, il est donc important de les identifier.

- Est-ce que les jugements (inférences) serviront à sélectionner certaines étudiantes ou certains étudiants? (2^e année? Spécialisation? Emploi? Etc.)
- Est-ce que les jugements (inférences) serviront à l'attribution d'un diplôme ou d'un titre?
- Est-ce que les jugements (inférences) serviront à la promotion ou la sélection de la personne enseignante?
- Et de nombreuses autres questions...

Les décisions et les communications devraient tenir compte des fonctions présentes. Dans l'ApC, la fonction de régulation des apprentissages est importante, elle devrait donc être mise de l'avant.

25

25



26



Les fardeaux de l'EdA de qualité

« La preuve dans un argumentaire est une série de revendications ou de propositions logiquement connectées qui sont appuyées par des données dans le cadre d'un mandat (Toulmin, 1958). Les revendications en matière d'évaluation sont relatives à la compétence des élèves. (Mislevy, 2009, p. 97) »

Mislevy propose donc de voir la démarche d'évaluation comme une chaîne où le maillon le plus faible limite la qualité des inférences.



Les fardeaux de l'EdA de qualité

La validation est le processus de cueillette de preuves (ou d'arguments) afin de soutenir la qualité des décisions prises.

La qualité des décisions ne peut être présumée, elle doit être démontrée.

ÉVALUER LE RAISONNEMENT CLINIQUE

Journées du CEEIADE 22 et 23 mars 2023

30

30

Évaluer le raisonnement clinique

« Celui-ci peut être défini comme l'ensemble des processus de pensée et de prises de décisions associés à la pratique clinique, qui permettent au clinicien de réaliser la meilleure action possible dans un contexte donné.»

(Higgs & Jones, 2008)

Journées du CEEIADE 22 et 23 mars 2023

31

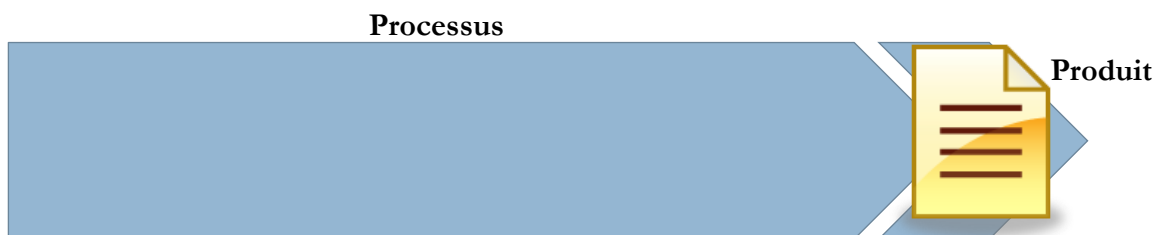
31

Évaluer le raisonnement clinique

« L'évaluation du raisonnement clinique est généralement considérée comme une tâche complexe (Charlin, Bordage, & van der Vleuten, 2003; Gruppen & Frohna, 2002; van der Vleuten, Norman, & Schuwirth, 2008). Cette complexité tient en partie au fait que le raisonnement clinique est imbriqué aux connaissances et aux habiletés en tant que composante de la compétence clinique (Charlin et al., 2003). »

(Dumas, 2016)

Évaluer le raisonnement clinique



Évaluer le raisonnement clinique

Les tâches les plus cohérentes avec l'évaluation du raisonnement clinique

- 1) Les QCM
- 2) Questions de type ouvertes (*Modified essay question*)
- 3) Test de concordance des scripts (TCS)
- 4) Problèmes de raisonnement clinique (*Clinical reasoning problem*)
- 5) Les examens cliniques objectifs structurés (ECOS)
- 6) L'évaluation du raisonnement clinique en contexte de stage

ÉVALUER DANS LE CADRE DE STAGE

- Rappelons les trois piliers de l'évaluation dans l'ApC.

Évaluer dans le cadre de stage

La mobilisation des connaissances, et non pas seulement leur mémorisation.

Les stages offrent souvent des occasions de présenter aux stagiaires des problèmes nouveaux et authentiques.

Ce sont donc des contextes où les tâches proposées sont souvent complexes.

Évaluer dans le cadre de stage

La trajectoire de développement.

Selon la logique de la trajectoire de développement, les stages devraient être organisés pour permettre une complexification des tâches.

La complexité des tâches peut être modifiée en jouant sur différentes variables:

- L'autonomie du stagiaire
- Le nombre d'univers à gérer simultanément
- La complexité des processus cognitifs

Évaluer dans le cadre de stage

La priorité de la fonction d'aide à l'apprentissage de l'évaluation.

Dans une approche par compétence, chaque tâche est à la fois une source d'apprentissage et d'évaluation.

Dans le stage, cela veut dire une rétroaction fréquente et cohérente tout au long du stage.

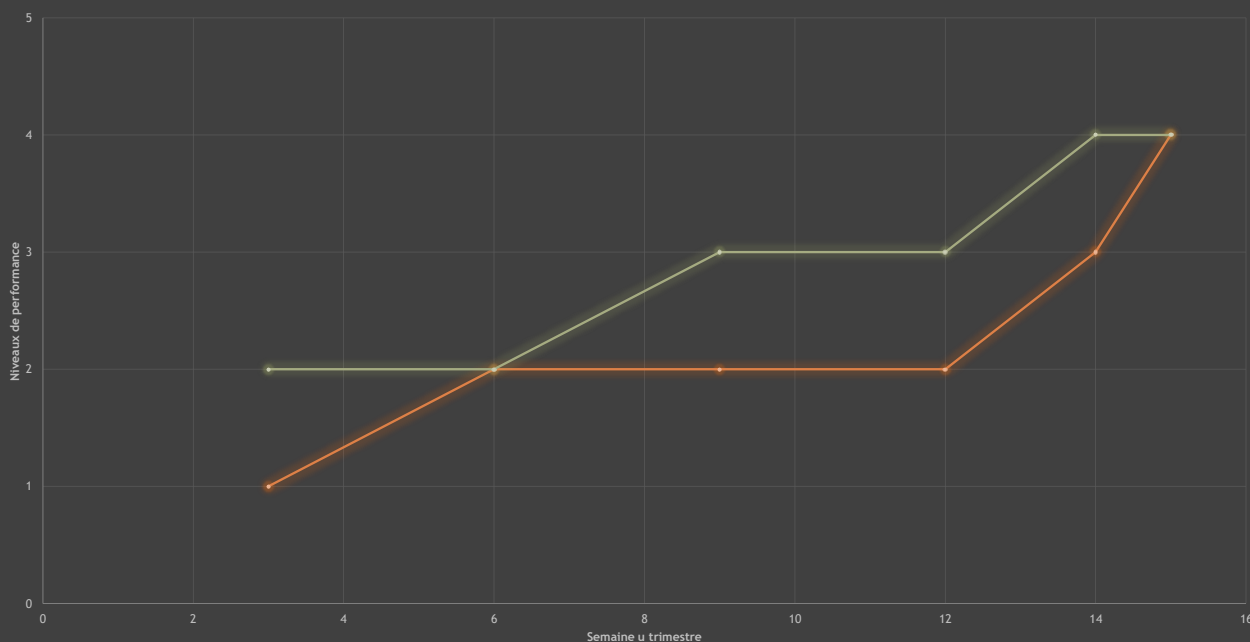
Journées du CEELADE 22 et 23 mars 2023



38

Trajectoire du développement pour la compétence 3

— Léo-Paul — Niveau d'attente



39



Évaluer dans le cadre de stage

La qualité de l'évaluation dans le cadre de stage devrait permettre de déterminer le niveau de compétence de chaque stagiaire sans influence du milieu de stage, de la personne mentor ou des événements spécifiques.

Le problème, c'est que souvent, dans les stages, chaque personne stagiaire est supervisée par une seule personne à qui l'on demande de faire à la fois la tâche de supervision et d'évaluation.

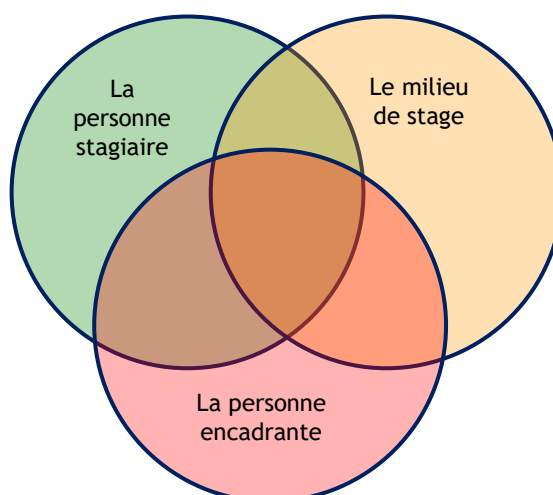
Journées du CEELADE 22 et 23 mars 2023

40

40



Évaluer dans le cadre de stage



Journées du CEELADE 22 et 23 mars 2023

41

41



Évaluer dans le cadre de stage

Parfois, aussi, la personne superviseure doit faire un bilan en s'appropriant le « langage » et les outils fournis par le programme de formation, ce qui lui demande de comprendre un univers différent dans lequel elle navigue tous les jours.

Pour compenser, on demande parfois à une personne liée directement au programme de remplir les outils, alors que cette personne n'a pas beaucoup observé la personne stagiaire en action.

Donc, comment s'outiller pour améliorer la qualité des inférences.

Journées du CEEIADE 22 et 23 mars 2023

42

42

LES OUTILS D'ÉVALUATION COHÉRENTS AVEC L'APC ET LE CONTEXTE DES STAGES

- Dans l'ApC, en plus des tâches, le principal outil d'évaluation est l'**échelle descriptive**
- En stage, son utilisation est parfois difficile cependant.

Journées du CEEIADE 22 et 23 mars 2023

43

43

C1 : Anticiper et mettre en place une organisation du site d'anesthésie en fonction du patient, du type d'intervention et du type d'anesthésie				
Entourer la lettre correspondant au critère atteint	D	C	B	A
	Les conditions d'exercice ne permettent pas de travailler en sécurité (manque ou délai trop long).	L'organisation est conforme et respecte les éléments de sécurité.	L'organisation est rapide, efficace et ergonomique. Les situations d'urgence potentielles sont connues ou repérées et expliquées.	L'organisation est rapide, efficace en toute situation (complexe, urgente ...).
Points forts :				
Axes de progression :				
Éléments justifiant l'attribution d'un critère D :				
Compétence 2: Analyser la situation, anticiper les risques associés à des caractéristiques du patient et de l'intervention et ajuster la prise en charge anesthésique				
Entourer la lettre correspondant au critère atteint	D	C	B	A
Analyse	Les éléments inhérents à la situation ne sont pas pris en compte.	L'analyse met en lien les éléments recueillis avec les connaissances théoriques et permet d'identifier les risques.	L'analyse permet la mise en place d'actions adaptées en situation simple. La situation clinique est comprise.	L'analyse permet de prévoir les actions préventives et/ou correctives adaptées et priorisées en situation complexe.
Entourer la lettre correspondant au critère atteint	D	C	B	A
Anticipation	Peu ou pas d'anticipation.	Anticipation des événements majeurs prévisibles en situation courante.	Anticipation des événements prévisibles dans une priorisation adaptée.	Anticipation des événements prévisibles dans une priorisation adaptée dans des situations complexes.
Points forts :				
Axes de progression :				
Éléments justifiant l'attribution d'un critère D :				

Critère

Niveau

Descripteur

1 critère x 1 niveau = 1 descripteur

Journées du CEEIADE 22 et 23 mars 2023 Document CEEIADE version 1 validée par le CA le 29/06/12

44



Recueillir les données

Compétence 5 - Évaluer la progression des apprentissages et le degré d'acquisition des compétences des élèves pour les contenus à faire apprendre

Indication d'un niveau de maîtrise de compétence	Maîtrise marquée (5)	Maîtrise assurée (4)	Maîtrise acceptable (3)	Maîtrise peu développée (2)	Maîtrise très peu développée (1)	w
Critère 5.1a Détecter les forces et les difficultés des élèves Démarche (planification spécifique)	Les tâches suggérées permettront de recueillir de l'information sur les forces et les faiblesses de chaque élève. Le choix et l'organisation des tâches sont optimaux et démontrent une maîtrise marquée de la planification spécifique.	La planification spécifique est complète. Les tâches suggérées permettront de recueillir de l'information sur les forces et les faiblesses de chaque élève.	La planification spécifique est complète. Les tâches suggérées permettront de recueillir de l'information sur les forces et les faiblesses de chaque élève.	Les tâches suggérées sont intéressantes. Il subsiste un doute si elles permettent de recueillir de l'information sur les forces et les faiblesses de chaque élève.	Certaines tâches ne sont pas appropriées pour les élèves ciblés par l'évaluation des apprentissages.	1
Critère 5.3a Contribuer à la préparation du matériel d'évaluation Démarche (prise d'information)	La création et l'administration de chacune des tâches sont décrites en détail. Leur arrimage est aussi expliqué en détail.	Un grand souci est accordé à ce que la création et l'administration des tâches d'évaluation permettent de s'assurer de la qualité du déroulement de l'évaluation.	La création des tâches et l'administration des tâches sont expliquées. Elles tiennent compte des précautions habituelles.	La préparation du matériel, la distribution des consignes et l'administration des tâches sont sommairement décrites.	La préparation du matériel d'évaluation n'est pas décrite de façon explicite OU elle va à l'encontre de pratiques minimales admises.	1
Critère 5.3b Contribuer à la préparation du matériel d'évaluation Démarche (interprétation)	Tous les choix d'outils sont expliqués et ces explications permettent d'appuyer tout le dispositif d'EdA.	Les outils permettent d'évaluer la progression et l'acquisition des compétences chez les élèves.	Les outils d'interprétation choisis et utilisés tiennent compte de la complexité des compétences, mais pourront difficilement en faire une évaluation satisfaisante chez chaque élève.	Les choix d'outils sont très peu recherchés ou très peu expliqués. Il est peu probable que ces choix permettent de faire l'évaluation de compétence pour chaque élève.	Les choix d'outils servant à l'EdA ne sont pas justifiés OU les justifications sont peu crédibles.	1
Critère 5.1b Détecter les forces et les difficultés des élèves Démarche (jugement et décision)	La démarche permet de porter un jugement solidement appuyé et qui tient compte des différents profils de chaque élève. Les modalités de communication	La démarche permet de porter un jugement bien appuyé sur les forces et les difficultés de chaque élève.	La planification prévoit le passage des informations recueillies au jugement. Cependant, il s'agit d'un processus simple.	Les inférences sont faites en compilant simplement les notes cumulées, et ce, de façon simplement mathématique.	Les inférences sont faites sans tenir compte réellement des informations recueillies tout au long de l'année.	1

Journées du CEEIADE 22 et 23 mars 2023

45

45

Recueillir les données

Une fois utilisée, l'échelle descriptive permet de recueillir différentes observations sur la performance d'une étudiante ou d'un étudiant et aussi de son groupe. Cela peut ressembler à ceci :

Groupe	ID.pers	C05.1	C05.2	C05.3	C05.4	C05.5	C05.6	C05.7	C05.8	C05.9	C02.1	C02.2	C02.3	C02.8	H00.a	H00.b
		2	3	1	2	3	2	3	1	1	2	2	1	2	1	1
A	A.001	5	5	5	5	5	5	5	5	5	4	4	4	4	3	3
A	A.002	3	2	3	3	3	3	2	2	4	4	3	3	3	3	3
A	A.003	4	3	4	4	4	4	4	3	3	4	4	4	4	4	4
A	A.004	4	3	4	4	3	4	3	3	3	4	4	4	4	3	2
A	A.005	3	1	3	2	2	2	2	2	2	3	2	3	2	2	2
A	A.006	2	2	3	3	3	3	2	2	2	3	2	3	3	3	2
A	A.007	4	4	4	4	4	4	4	4	3	4	3	4	4	3	3
A	A.008	4	3	3	3	2	3	4	3	4	4	3	4	4	4	4
A	A.009	4	4	4	5	4	5	5	4	4	4	4	4	4	4	5
A	A.010	4	4	4	5	4	4	3	4	4	4	5	4	4	5	3
A	A.011	4	3	4	4	4	3	3	3	4	4	4	4	4	4	3
A	A.012	2	2	3	3	3	3	2	3	3	4	2	4	4	3	3
A	A.013	3	3	3	3	3	4	4	4	3	3	3	4	4	3	3
A	A.014	3	3	4	4	3	4	3	4	4	4	4	4	4	3	3
A	A.015	4	3	3	4	2	4	3	3	4	4	3	4	4	3	3
A	A.016	3	4	4	3	3	4	3	4	4	3	4	4	4	4	4
A	A.017	3	4	4	3	4	2	3	3	4	4	3	3	3	4	4
A	A.018	4	5	4	4	3	4	4	5	4	4	4	4	4	4	4
A	A.019	4	4	4	4	3	3	3	4	4	4	4	4	4	4	4

Journées du CEELADE 22 et 23 mars 2023

46

46

Pourquoi l'échelle descriptive?

En s'intéressant aux qualités des performances des étudiantes et des étudiants, l'échelle descriptive permet de recueillir des données sur les performances, même si elles diffèrent entre elles.

Cependant, ce n'est pas une panacée. Pour s'accomplir correctement de la tâche, il faut s'assurer d'être attentif avant, pendant et après son utilisation.

Avant, soit lors de sa création.

Pendant, soit durant son utilisation.

Après, une fois les observations faites.

Journées du CEELADE 22 et 23 mars 2023

47

47



L'approche critériée et les échelles descriptives

La création d'une échelle descriptive est en soi une tâche complexe. Plusieurs auteurs proposent des façons d'élaborer ces échelles en proposant une procédure permettant d'identifier et de structurer les critères et de déterminer le nombre de niveaux. (Leroux, 2015; Côté, Tardif et Munn, 2011; Côté, 2014, Stevens et Levi, 2005)

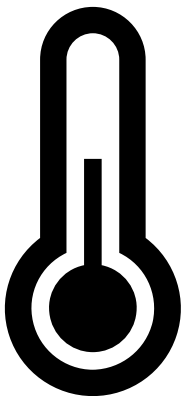
Selon Laurier (2005), trois éléments permettent l'obtention d'une grille de qualité.

Soit...

- la pertinence du critère
- l'observabilité du critère
- l'univocité du critère



Concevoir une échelle



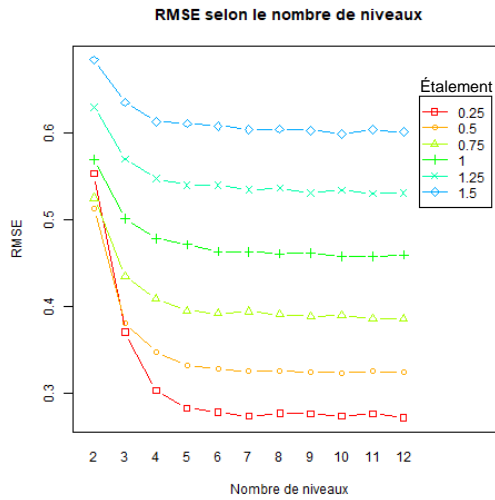
Quand on conçoit une échelle descriptive, il est important de couvrir l'ensemble des niveaux potentiellement observables, et ce, tout au long de la trajectoire de développement.

Chaque niveau devrait décrire un palier dans la trajectoire de développement.



Concevoir une échelle

Quel est le nombre optimal de niveaux?

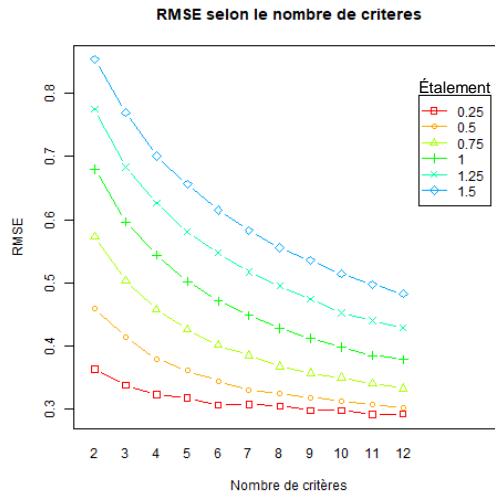


51



Concevoir une échelle

Quel est le nombre optimal de critères?



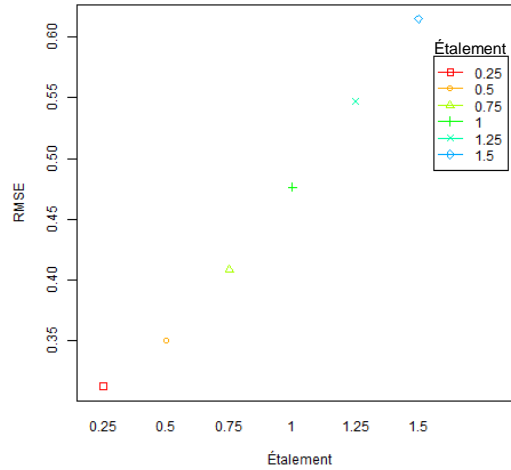
52



Concevoir une échelle

Quel est l'étalement optimal?

RMSE selon la valeur d'étalement



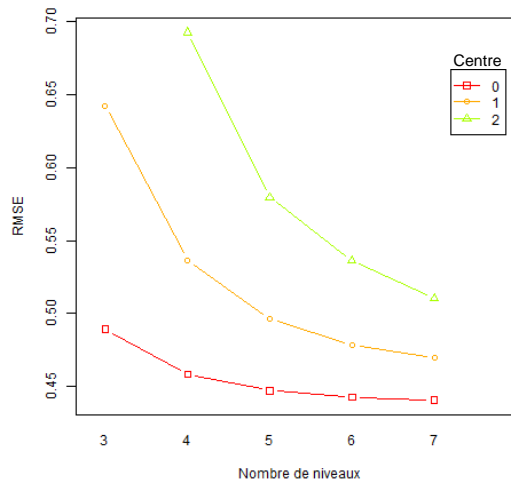
53



Concevoir une échelle

Est-il mieux d'utiliser des échelles centrées ou décentrées?

RMSE selon le nombre de niveaux



54



Pendant l'utilisation de l'échelle

Les échelles aident à la fidélité des observations, mais seulement si elles sont bien utilisées.

Concrètement, cela veut dire former les utilisatrices et les utilisateurs avec de réels cas typiques des différents niveaux.

Dans le contexte de stage, il faut aussi s'assurer que les comportements recherchés puissent être observés directement par la personne qui utilise l'échelle descriptive.



Après, caractériser notre échelle

Si l'échelle est bien conçue et bien utilisée, elle permettra de recueillir beaucoup d'observations.

L'utilisation **de l'analyse d'items** s'appuyant sur des données reproduisant le contexte de l'utilisation de tâches complexes et d'échelles descriptives. Les outils les plus pertinents seront présentés ainsi que les limites à respecter afin de les utiliser judicieusement.

Cette analyse confirmera que l'échelle remplit sa fonction correctement. Elle devrait aussi nous permettre d'améliorer l'échelle avant sa prochaine utilisation.



Après, caractériser notre échelle

Nous avons retenu, pour aujourd'hui, deux groupes d'outils:

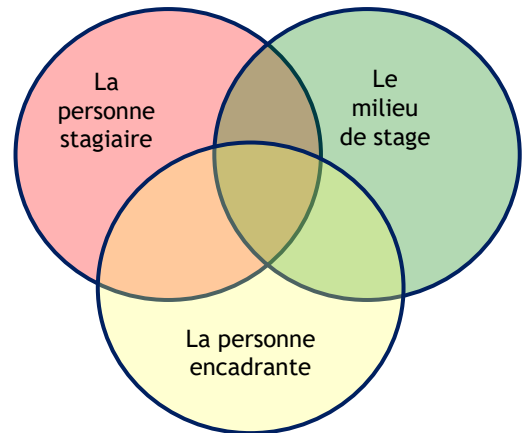
Ceux de la théorie classique des tests (TCT), dont

- L'indice de difficulté (P)
- L'indice de discrimination (D)
- Cohérence interne

Ceux de la théorie des réponses aux items (TRI)

- Habileté (Théta)
- Difficulté (a)
- Discrimination (b)
- Pseudo-chance (c)

Journées du CEEIAD 22 et 23 mars 2023



57

57



Après, caractériser notre échelle

La présente communication montre l'utilisation **de l'analyse d'items** s'appuyant sur des données réelles reproduisant le contexte de l'utilisation de tâches complexes et d'échelles descriptives. Les outils les plus pertinents seront présentés ainsi que les limites à respecter afin de les utiliser judicieusement.

Auger (2000) souligne que bien que cela soit rarement fait, il n'y a pas de raison que les outils de la mesure ne soient pas utilisés dans une approche utilisant des tâches complexes.

Journées du CEEIAD 22 et 23 mars 2023

58

58

Améliorer notre échelle

Si on a deux groupes ou deux promotions, il est possible de comparer les valeurs des groupes.

		56 51	56 41	56 56	55 37	55 39	55 45	56 42	56 46	56 47
A										
Moyenne	3,49	3,16	3,62	3,22	2,95	3,29	3,11	3,25	3,29	
difficulté1	0,70	0,63	0,72	0,64	0,59	0,66	0,62	0,65	0,66	
difficulté2	0,91	0,73	1,00	0,67	0,71	0,82	0,75	0,82	0,84	
point.biserial	0,68	0,66	0,52	0,79	0,69	0,71	0,65	0,65	0,52	
polyserial	0,78	0,71	0,62	0,84	0,74	0,76	0,69	0,71	0,57	
		51 47	51 42	51 50	51 33	51 36	51 40	51 39	51 49	51 41
B										
Moyenne	3,53	3,41	3,76	3,16	3,14	3,31	3,43	3,59	3,43	
difficulté1	0,71	0,68	0,75	0,63	0,63	0,66	0,69	0,72	0,69	
difficulté2	0,92	0,82	0,98	0,65	0,71	0,78	0,76	0,96	0,80	
point.biserial	0,58	0,64	0,71	0,81	0,82	0,76	0,75	0,65	0,60	
polyserial	0,66	0,72	0,90	0,89	0,88	0,83	0,85	0,75	0,68	

Journées du CEEIAD 22 et 23 mars 2023

59

59

Après, caractériser notre échelle

La théorie des réponses aux items permet à la fois de caractériser les personnes et les items. Dans le modèle à trois paramètres, chaque item est modélisé par 3 paramètres.

- La pseudo-chance (c)
- La difficulté (a)
- La discrimination (b)
- Elle permet aussi de déterminer l'habileté de chaque personne étudiante (Théta)

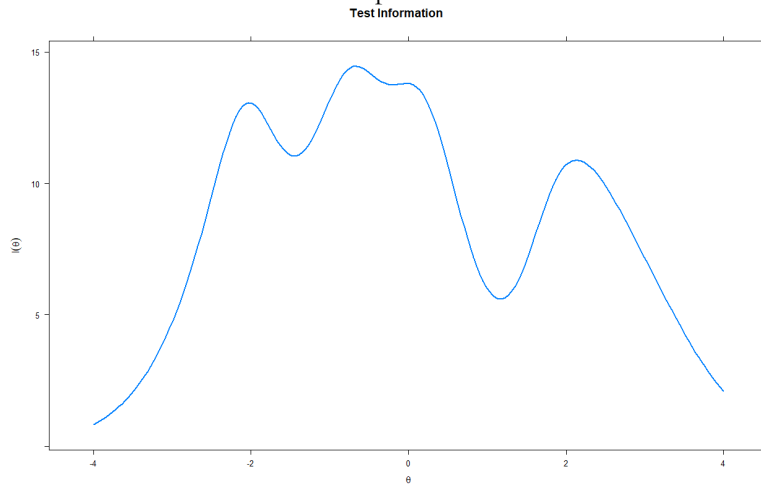
Journées du CEEIAD 22 et 23 mars 2023

60

60

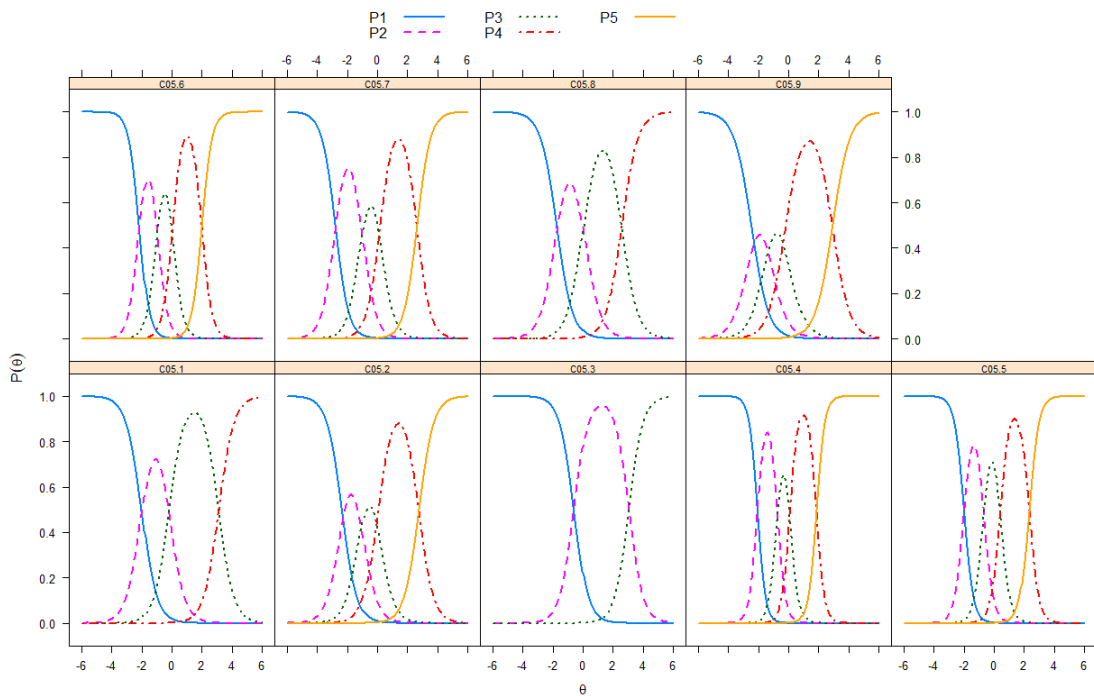
Les outils de la TRI

Des éléments intéressants sur la tâche complète.



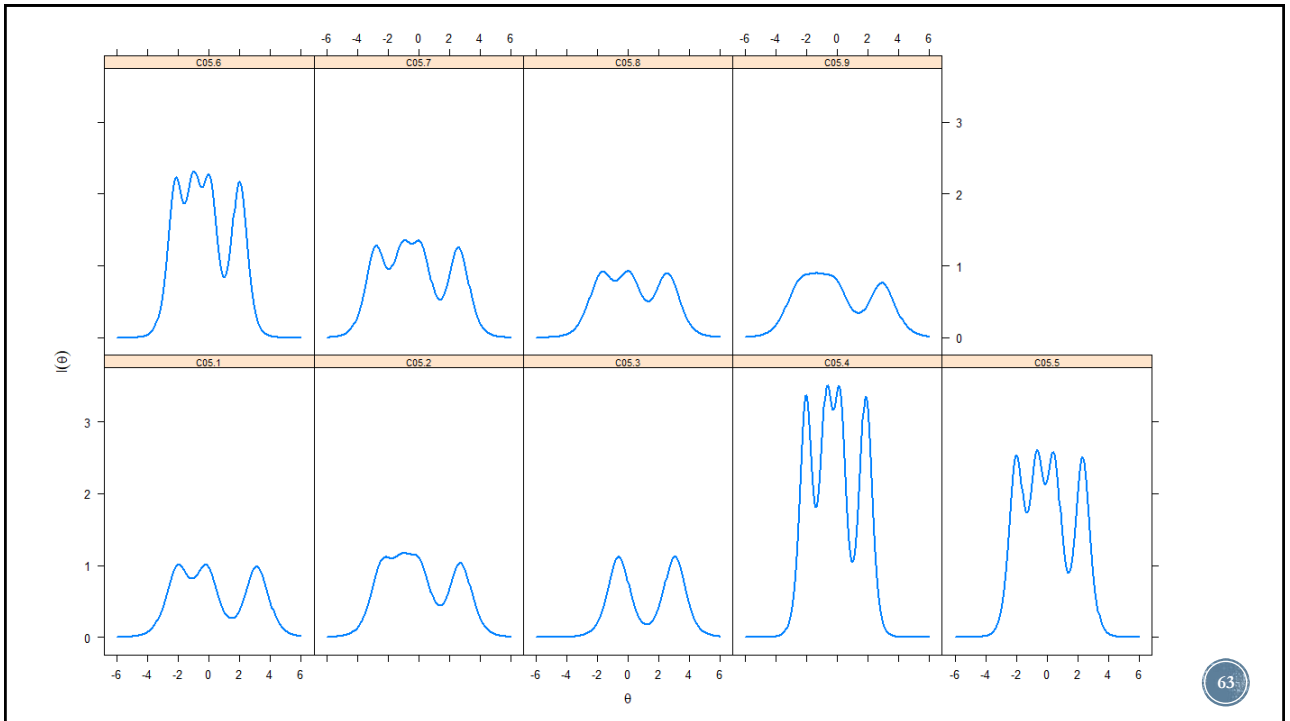
61

61



62

62



63

Utiliser la moyenne pondérée

Traditionnellement, on utilisait la somme (pondérée ou non) pour faire la compilation des critères dans une même tâche. Nous suggérons la moyenne pondérée à la place.

B	C	D	E	F	G	H	I	J	K	L	M
ID.pers	C05.1	C05.2	C05.3	C05.4	C05.5	C05.6	C05.7	C05.8	C05.9	Moyenne pondérée	fscores
A.001	2	3	1	2	3	2	3	1	1	5,0	3,44
A.002	5	5	5	5	5	5	5	5	5	2,8	-0,71
A.005	3	2	3	3	3	3	2	2	4	2,1	-1,83
A.006	3	2	3	3	3	3	2	2	2	2,4	-1,00
A.007	4	4	4	4	4	4	4	4	3	3,9	0,88
A.008	4	3	3	3	2	3	4	3	4	3,2	-0,37
A.009	4	4	4	5	4	5	5	4	4	4,3	2,02
A.010	4	4	4	5	4	4	3	4	4	4,0	1,24
A.011	4	3	4	4	4	3	3	3	4	3,6	0,23
A.012	2	2	3	3	3	3	2	3	3	2,7	-0,79
A.013	3	3	3	3	3	4	4	4	3	3,3	-0,12
A.014	3	3	4	4	3	4	3	4	4	3,6	0,21
A.015	4	3	3	4	2	4	3	3	4	3,3	-0,11

	Maîtrise marquée (5)		Maîtrise assurée (4)		Maîtrise acceptable (3)		Maîtrise peu développée (2)		Maîtrise très peu développée (1)		
A.005	3	1	3	2	2	2	2	2	2	2,1	-1,83
A.006	2	2	3	3	3	3	2	2	2	2,4	-1,00
A.007	4	4	4	4	4	4	4	4	3	3,9	0,88
A.008	4	3	3	3	2	3	4	3	4	3,2	-0,37
A.009	4	4	4	5	4	5	5	4	4	4,3	2,02
A.010	4	4	4	5	4	4	3	4	4	4,0	1,24
A.011	4	3	4	4	4	3	3	3	4	3,6	0,23
A.012	2	2	3	3	3	3	2	3	3	2,7	-0,79
A.013	3	3	3	3	3	4	4	4	3	3,3	-0,12
A.014	3	3	4	4	3	4	3	4	4	3,6	0,21
A.015	4	3	3	4	2	4	3	3	4	3,3	-0,11

64

Conclusion

La qualité de l'évaluation est tributaire du maillon le plus faible, soit le maillon sur lequel nous avons le moins de preuves.

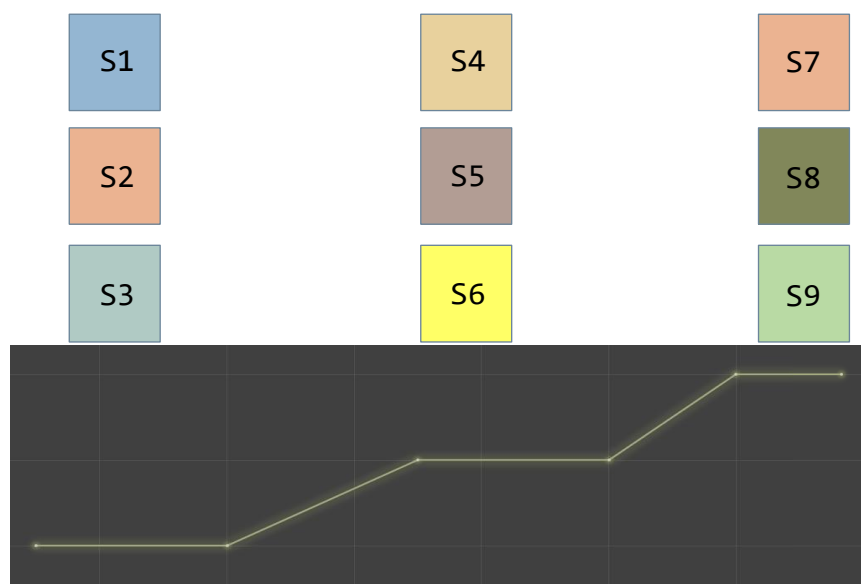
C'est un processus itératif, ce qui est important c'est que l'année X+1 soit plus résistante que l'année X.

Améliorer la qualité de l'évaluation amène nécessairement une bonification de l'alignement pédagogique.

65

65

Conclusion



66

66

Conclusion

Un des premiers maillons à travailler est une échelle descriptive, puisqu'elle est une autre façon de représenter la trajectoire de développement et rend concrète l'observation du développement de la compétence.

Compétence 1 : Résoudre des problèmes 2^e cycle Compétence 2 : Écrire des textes variés 3^e cycle

Échelon 1 Échelon 2 Échelon 3 Échelon 4 Échelon 5 Échelon 6 Échelon 7 Échelon 8 Échelon 9 Échelon 10

Deviser une situation relativement simple, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Deviser une situation, l'élève exprime généralement en quel(s) objet(s) lui paraît le plus problématique pour lui et pour son environnement immédiat. Il prend appui sur son vécu pour tenter de résoudre le problème. Il émet des hypothèses et tente de les vérifier. Il utilise des connaissances et des compétences acquises au cours de sa scolarité. Il s'efforce de trouver une solution, même si elle n'est pas optimale. Il est capable de justifier sa solution.

Institut Villebon - Georges Charpak



Questions



Institut Villebon - Georges Charpak



Références

- Aubin, A.-S. (2015). Caractéristiques de l'évaluation des apprentissages dans une faculté de génie.
- Auger, R. 1949-, Séguin, S. P. 1944-, Nézet-Séguin, C. 1945-, Séguin, S. P. 1944- et Nézet-Séguin, C. 1945-. (2000). Formation de base en évaluation des apprentissages. Éditions Logiques.
- Baker, J. G., Rounds, J. B. et Zevon, M. A. (2000). A Comparison of Graded Response and Rasch Partial Credit Models with Subjective Well-Being. *Journal of Educational and Behavioral Statistics*, 25(3), 253.
- Côté, F. 1964-. (2014). Construire des grilles d'évaluation descriptives au collégial : guide d'élaboration et exemples de grille. Presses de l'Université du Québec.
- Fournier, J.-P., Thiercelin, D., Pulcini, C., Alunni-Perret, V., Gilbert, E., Minguet, J.-M. et Bertrand, F. (2006). Évaluation du raisonnement clinique en médecine d'urgence : les tests de concordance des scripts décèlent mieux l'expérience clinique que les questions à choix multiples à contexte riche. *Pédagogie Médicale*, 7(1), 20-30.
- Laurier, M. (Michel D.), Morissette, Dominique, Tousignant, Robert, Tousignant, Robert. et Morissette, D. 1933-. (2005). Les principes de la mesure et de l'évaluation des apprentissages (3e éd.). G. Morin.
- Laveault, D. et Grégoire, J. (2014). Introduction aux théories des tests en psychologie et en sciences de l'éducation. De Boeck Supérieur.
- Leroux, J. L., Groupe de recherche en évaluation des apprentissages au collégial, Association québécoise de pédagogie collégiale (2015). Évaluer les compétences au collégial et à l'université : un guide pratique. Association québécoise de pédagogie collégiale (AQPC).
- Popham, W. J. (2000). *Modern Educational Measurement: Practical Guidelines for Educational Leaders*. Allyn and Bacon.
- Québec (Province) et Ministère de l'éducation. (2003). Politique d'évaluation des apprentissages: formation générale des jeunes, formation générale des adultes, formation professionnelle. Éducation Québec.
- RStudio Team. (2019). RStudio: Integrated Development Environment for R. RStudio, Inc. <http://www.rstudio.com/Salkind>, N. J. (2012). *Tests & Measurement for People Who (Think They) Hate Tests & Measurement*. SAGE Publications.
- Stevens, D. D. et Levi, A. J. (2013). *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning*. Stylus Publishing, LLC.
- Tardif, J., Fortier, G. et Préfontaine, C. (2006). *L'évaluation des compétences: documenter le parcours de développement*. Chenelière Education.

