

Description du processus d'élaboration d'une grille d'évaluation spécifique à une situation d'apprentissage et d'évaluation en contexte de modification des contenus d'apprentissage auprès d'élèves du secondaire présentant un trouble du spectre de l'autisme

Description of the process of developing a rubric specific to a learning and evaluation situation in the context of modifying learning content for secondary school students with autism spectrum disorder

Descrição do processo de elaboração de uma rubrica de avaliação específica para uma situação de aprendizagem e avaliação no contexto de modificação dos conteúdos de aprendizagem para alunos do ensino secundário com transtorno do espectro do autismo

Marie-Aimée Lamarche et Micheline-Joanne Durand

Volume 45, numéro 3, 2022

Réception : 20 décembre 2022

Version finale : 19 juillet 2023

Acceptation : 26 août 2023

URI : <https://id.erudit.org/iderudit/1107235ar>

DOI : <https://doi.org/10.7202/1107235ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Lamarche, M.-A. & Durand, M.-J. (2022). Description du processus d'élaboration d'une grille d'évaluation spécifique à une situation d'apprentissage et d'évaluation en contexte de modification des contenus d'apprentissage auprès d'élèves du secondaire présentant un trouble du spectre de l'autisme. *Mesure et évaluation en éducation*, 45(3), 95–133. <https://doi.org/10.7202/1107235ar>

Résumé de l'article

L'élaboration d'une grille d'évaluation est une tâche complexe qui exige une bonne connaissance de l'objet d'évaluation et du niveau d'apprentissage des apprenants pour lesquels elle est destinée. Parmi les différents types de grilles d'évaluation qui sont utilisées en milieu scolaire, la grille descriptive critériée ainsi que la grille descriptive analytique s'avèrent des choix particulièrement appropriés pour guider le jugement des évaluateurs. En suivant la méthodologie de la recherche-développement, les concepteurs se sont engagés dans un processus réflexif qui, après de nombreux aller-retour dans la formulation des critères, des échelons et des descripteurs, a mené à l'élaboration, à la validation et à la mise à l'essai de trois versions d'une grille descriptive analytique. Les résultats permettent d'identifier les éléments à considérer dans la construction de cet outil pour des élèves du premier cycle du secondaire présentant un trouble du spectre de l'autisme (TSA) et une grande diversité de profils cognitifs.

© Marie-Aimée Lamarche et Micheline-Joanne Durand, 2023



Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

Description du processus d'élaboration d'une grille d'évaluation spécifique à une situation d'apprentissage et d'évaluation en contexte de modification des contenus d'apprentissage auprès d'élèves du secondaire présentant un trouble du spectre de l'autisme

Description of the process of developing a rubric specific to a learning and evaluation situation in the context of modifying learning content for secondary school students with autism spectrum disorder

Descrição do processo de elaboração de uma rubrica de avaliação específica para uma situação de aprendizagem e avaliação no contexto de modificação dos conteúdos de aprendizagem para alunos do ensino secundário com transtorno do espectro do autismo

Marie-Aimée Lamarche
ORCID: 0000-0001-8367-8279
Université de Montréal

Micheline-Joanne Durand
ORCID: 0000-0002-7540-0186
Université de Montréal



MOTS CLÉS: aménagement à l'évaluation, enseignement secondaire, grille descriptive analytique, recherche-développement, trouble du spectre de l'autisme

L'élaboration d'une grille d'évaluation est une tâche complexe qui exige une bonne connaissance de l'objet d'évaluation et du niveau d'apprentissage des apprenants pour lesquels elle est destinée. Parmi les différents types de grilles d'évaluation qui sont utilisées en milieu scolaire, la grille descriptive critériée ainsi que la grille descriptive analytique s'avèrent des choix particulièrement appropriés pour guider le jugement des évaluateurs. En suivant la méthodologie de la recherche-développement, les concepteurs se sont engagés dans un processus réflexif qui, après de nombreux aller-retour dans la formulation des critères, des échelons et des descripteurs, a mené à l'élaboration, à la validation et à la mise à l'essai de trois versions d'une grille descriptive analytique. Les résultats permettent d'identifier les éléments à considérer dans la construction de cet outil pour des élèves du premier cycle du secondaire présentant un trouble du spectre de l'autisme (TSA) et une grande diversité de profils cognitifs.

KEY WORDS: adapted assessment, autism spectrum disorder, development research, middle school, rubrics

The development of rubrics is a complex task that requires a good knowledge of the object of evaluation and the learning level of the students for whom it is intended. From the several types of rubrics used in school environments, the analytic rubrics and developmental rubrics prove to be especially appropriate choices to guide the judgment of the evaluators. By following the research and development methodology, the designers engaged in a reflective process which, after many round trips in the formulation of criteria, scale and descriptors, led to the development, validation and implementation testing of three prototypes of analytic rubrics. The results make it possible to identify the elements to be considered in the construction of this tool for pupils with autism spectrum disorder (ASD) in the first cycle of secondary school who present a great diversity of cognitive profiles.

PALAVRAS CHAVE: ensino secundário, investigação-desenvolvimento, organização da avaliação, rubrica analítica, transtorno do espectro do autismo

Elaborar uma rubrica é uma tarefa complexa que exige uma bom conhecimento do objeto de avaliação e do nível de aprendizagem dos alunos aos quais se destina. Entre os diferentes tipos rubricas que são utilizadas nas escolas, a rubrica descritiva referenciada a critérios, bem como a rubrica analítica, revelam-se escolhas particularmente apropriadas para orientar o juízo dos avaliadores. Seguindo a metodologia da investigação-desenvolvimento, os autores envolveram-se num processo reflexivo que, após inúmeras idas e vindas na formulação de critérios, níveis e descritores, levou à elaboração, à validação e à testagem de três versões de uma rubrica analítica. Os resultados permitem identificar os elementos a considerar na construção desta ferramenta para alunos do ensino secundário com transtorno do espectro do autismo (TEA) e uma ampla diversidade de perfis cognitivos.

Problématique

Depuis l'avènement du Programme de formation de l'école québécoise (PFEQ) (Ministère de l'Éducation du Québec (MEQ), 2001), l'approche par compétences a été adoptée et se fonde « sur une intégration des savoirs (théoriques et pratiques), des savoir-faire et des attitudes nécessaires à l'accomplissement, de façon satisfaisante, des tâches complexes, significatives pour l'élève et nécessaires à son adaptation ultérieure à la vie d'adulte » (Louis, 2004, p. 22). Dans un tel contexte, la mise en place de tâches complexes va au-delà de la mesure des connaissances restituées et vise plutôt l'observation de la construction de savoirs, soit la manifestation de stratégies cognitives et métacognitives (Gérard, 2010; Tardif, 1993). L'interprétation des traces d'apprentissage des élèves dans ce type de tâche demande une instrumentation spécifique basée sur des critères convenablement formulés. La grille d'évaluation descriptive analytique ou critériée (Berthiaume et al., 2011), nommée aussi *rubric* dans les écrits anglophones (Stevens & Levi, 2012), est définie comme un outil qui permet d'appuyer le jugement des évaluateurs concernant la qualité et la progression d'une performance, à l'aide de critères et de descripteurs précis pouvant être associés à une pondération ou à une notation en cote (Brookhart, 2018; Brookhart & Chen, 2015; Panadero & Jonsson, 2020; Yetiş, 2017). Par sa transparence et sa précision, c'est un instrument qui permet d'assurer une meilleure compréhension des différentes manifestations des apprentissages des élèves en les plaçant sur un continuum.

La grille d'évaluation descriptive intéresse de plus en plus les chercheurs et une diversité de sujets d'étude s'y rattachent. Certains auteurs se penchent sur la compréhension et sur l'usage de l'instrument (Baribeau, 2009, 2015; Chan & Ho, 2019; Dawson, 2017; Wiertz et al., 2020), alors que d'autres se concentrent sur les éléments qui en font un outil efficace (Brookhart & Chen, 2015; Durand & Trépanier, 2011). Dans le cadre du présent article, nous nous intéressons à la construction de grilles d'évaluation dans une perspective d'évaluation en soutien à l'apprentissage alors que les apprenants évoluent dans le secteur de l'adaptation scolaire et vivent avec des conditions cognitives particulières, soit un trouble du spectre de l'autisme

(TSA). Ce trouble se manifeste, entre autres, par un fonctionnement atypique de la pensée dans certaines circonstances. Par exemple, ces élèves peuvent éprouver des difficultés à manifester de l'empathie envers les autres (renvoi à la théorie de l'esprit); il peut leur être difficile de traiter des informations dans leur entièreté sans s'attarder aux détails (renvoi à la cohérence centrale) et ils peuvent éprouver des difficultés sur le plan organisationnel, ce qui les conduit à avoir besoin d'activités routinières (renvoi aux fonctions exécutives) (Frith, 2003; Vermeulen, 2014). À notre connaissance, la construction de grilles d'évaluation descriptives dans un tel contexte n'a encore jamais été documentée scientifiquement (Lamarche & Durand, 2021).

L'évaluation des élèves présentant un trouble du spectre de l'autisme

Dans le système scolaire québécois, des cadres d'évaluation proposent des critères d'évaluation qui ne conviennent pas toujours aux profils diversifiés des élèves présentant un TSA. Ces derniers ont des profils d'apprentissage et des profils cognitifs atypiques et ne conçoivent pas la vie de la même façon que les jeunes neurotypiques (Vermeulen, 2014). De ce fait, il devient difficile d'utiliser les mêmes critères d'évaluation pour juger de leurs apprentissages. Dans ce contexte, les enseignants peuvent utiliser des aménagements qui se situent à différents niveaux et qui influencent ou non leur jugement. Par exemple, ils peuvent utiliser des outils numériques qui soutiennent les processus mentaux des élèves et qui n'influencent pas le jugement, comme la reconnaissance vocale. Ils peuvent aussi simplifier ou retirer des critères d'évaluation ce qui, en revanche, influence directement le jugement évaluatif. Cette distinction est approfondie dans la section suivante. Selon Jung et Guskey (2011), les aménagements inscrits au plan d'intervention (PI) des élèves guident l'élaboration du design des instruments d'évaluation, c'est-à-dire des grilles d'évaluation. De plus, les auteurs précisent que l'interprétation de productions d'élèves ayant droit à des aménagements représente un défi supplémentaire pour l'enseignant, car ce dernier doit distinguer l'apprentissage de l'élève de son niveau d'autonomie dans la réalisation de la tâche. La grille d'évaluation descriptive analytique semble être ici un outil pertinent pour soutenir les apprentissages et intégrer les aménagements nécessaires au cheminement scolaire des élèves présentant un TSA.

Les aménagements apportés à la tâche

Plusieurs élèves vivant avec un TSA présentent des défis cognitifs significatifs qui nécessitent des aménagements importants dans les pratiques enseignantes, tant lors de l'apprentissage que de l'évaluation, notamment, la mise en place de dispositifs de différenciation pédagogique.

Pour désigner l'ensemble des dispositifs de différenciation, les auteurs anglophones (Calhoun et al., 2000 ; Fuchs & Fuchs, 2001 ; Salend & Duhaney, 2002 ; Scanlon & Baker, 2012) les termes *inclusive setting*, *adaptations* ou *test accommodations*. Dans les écrits scientifiques francophones, les auteurs (Denis et al., 2016 ; Meirieu, 2016 ; Paré, 2012 ; Poirier et al., 2017) utilisent des termes comme *accommodations*, *adaptations*, *aménagements* ou *pratiques différenciées*. Tous ces termes permettent d'adapter les pratiques dans une perspective d'intégration des élèves dans un contexte dit ordinaire. Par ailleurs, les documents ministériels du Québec (2014) identifient trois niveaux d'interventions différenciées : la flexibilité pédagogique (interventions communes et réfléchies en fonction de la diversité des élèves), les mesures d'adaptation (interventions qui ont pour but d'atténuer les barrières à l'apprentissage sans diminuer le niveau d'exigence) et les mesures de modification (interventions qui nécessitent un PI et qui ont pour but de changer les attentes et les exigences du PFEQ (Ministère de l'Éducation, du Loisir et du Sport (MELS), 2006)).

Dans le cadre de notre projet, les enseignantes ont été amenées à différencier leurs modalités d'évaluation en ce qui concerne les contenus d'apprentissage et à modifier leurs attentes. Nous avons retenu l'appellation aménagements majeurs lorsque le contenu enseigné et évalué se retrouve en deçà des exigences du programme de formation ciblé pour le groupe-âge. Dans le même ordre d'idée, lorsque les dispositifs de différenciation permettent d'atténuer les obstacles à l'apprentissage sans changer le niveau d'exigence, nous avons retenu l'appellation aménagements mineurs.

Les grilles d'évaluation sont difficiles à utiliser avec des élèves présentant un TSA, car les apprentissages réalisés en classe ne répondent pas aux critères d'évaluation prescrits par le ministère (Lamarche & Durand, 2022 ; MELS, 2011). La recherche entourant le développement de grilles d'évaluation des compétences scolaires pour ce type d'apprenants est très peu développée. Bien que différentes pratiques enseignantes et évaluatives aient été documentées (Poirier et al., 2017), l'interprétation de l'évaluation des compétences scolaires reste un défi de taille (Lamarche & Durand, 2021, 2022).

La recension des écrits

Les grilles d'évaluation descriptives analytiques ont fait l'objet de différentes recherches qui mettent en évidence les pratiques d'utilisation de l'outil ainsi que les avantages et les défis associés. Alors que plusieurs auteurs s'intéressent directement à cet outil au niveau collégial (Chaumont, 2015; Côté, 2014; Leroux, 2010; Mastracci, 2011; Vincent, 2017) et universitaire (Berthiaume et al., 2011; Martin et al., 2016), des recherches menées au primaire et au secondaire abordent l'outil à travers les pratiques évaluatives des enseignants. En effet, Ramoo et Durand (2016) ont documenté les pratiques évaluatives d'enseignants travaillant en 6^e année du primaire. Leurs résultats indiquent que les enseignants utilisaient davantage les grilles d'évaluation à échelles uniformes numériques et à échelles descriptives lors des activités d'écriture que lors des activités de mathématiques et de lecture. Dans son étude, Brind'Amour (2018) documente les pratiques évaluatives d'enseignants de première année du primaire pour la compétence *Lire des textes variés*. Bien que la grille d'évaluation descriptive fasse partie des instruments pouvant être utilisés, les résultats démontrent que les enseignants utilisent plutôt la grille d'observation. Deudelin et al. (2007) ont de leur côté documenté les pratiques d'évaluation formatives de 13 enseignants du primaire. Leurs résultats indiquent que les enseignants apprécient utiliser des grilles d'évaluation maison pour soutenir l'autoévaluation, le travail de coopération et le processus de résolution de problèmes en mathématiques. L'utilisation de cet instrument vise à faire participer l'élève à son évaluation et, surtout, à son apprentissage. Au secondaire, Di Lalla (2017) et Baribeau (2009, 2015) ont observé les pratiques d'enseignants à la fin d'un cycle. Dans ces deux contextes, les enseignants utilisaient les grilles d'évaluation fournies par le gouvernement lors des épreuves ministérielles afin de documenter leur jugement. Ces études montrent que les grilles sont utilisées dans les classes québécoises, mais aucune ne se penche sur leur élaboration et n'aborde le secteur de l'adaptation scolaire où les aménagements sont nécessaires, alors que cet instrument pourrait y être approprié.

Tout d'abord, Goodrich (1996) soutient que la grille d'évaluation est un instrument flexible qui permet de prendre en considération les profils hétérogènes des élèves d'une même classe. En effet, selon les capacités des élèves, il est possible d'ajuster les attentes dans la grille en ajoutant ou en retirant des niveaux dans les échelons (Goodrich, 1996). De plus, différentes recensions des écrits ont clairement montré que ce type de

grille d'évaluation soutient l'apprentissage des élèves, leur développement de l'autorégulation ainsi que celui des enseignants, et peut influencer positivement la motivation des élèves (Brookhart & Chen, 2015 ; Jonsson & Svingby, 2007 ; Panadero & Jonsson, 2013). Ces éléments positifs découlent, entre autres, de la transparence qui a été documentée par Berthiaume et al. (2011) en contexte universitaire lors d'un cours d'histoire. L'auteur précise que la grille permet effectivement aux étudiants de mieux comprendre la tâche à réaliser, mais aussi aux enseignants de mieux planifier leur enseignement.

Par ailleurs, Martin et al. (2016) et Wiertz et al. (2020) mettent en exergue l'aspect chronophage de la construction d'une grille d'évaluation descriptive (il n'y a donc pas de gain de temps avec cet instrument) ainsi que l'importance d'avoir une compréhension univoque de son contenu. Pour ces auteurs, la grille est un instrument complexe qui n'est pas toujours facile à utiliser, compte tenu du lexique qu'elle contient. Wiertz et al. (2020) précisent que, pour surmonter cet obstacle, il est bénéfique de construire la grille en équipe et de confronter ses idées. C'est aussi ce qu'ont constaté Dolz-Mestre et Tobola Couchepin (2015) qui ont inclus des élèves dans la construction d'une grille.

Dans leur récente recension des écrits, Panadero et Jonsson (2020) ont exploré les critiques à propos des grilles d'évaluation en général. Les résultats de leur recension ont permis de relever six thèmes¹ qui, selon les auteurs, permettent de constater que les preuves empiriques à la base de ces critiques ne sont pas solides et s'appuient, pour la plupart, sur des événements anecdotiques. Ils relèvent également une mécompréhension du concept par les auteurs des différentes recherches constituant le corpus recensé. Ils concluent que la grille d'évaluation peut être utilisée de maintes façons et que la conceptualisation de l'outil peut être variable. Néanmoins, ils précisent que nous « pourrions optimiser la conception et la mise en œuvre des grilles grâce à la recherche scientifique portant sur les bénéfices et les limites de l'outil² » (Panadero & Jonsson, 2020, p. 17).

-
1. « The themes have been named: a) Standardization and narrowing the curriculum, b) Instrumentalism and “criteria compliance”, c) Simple implementations don't work, d) Limitations of criteria, e) Context dependence, and f) Miscellanea. » (Panadero et Jonsson, 2020, p. 14)
 2. Traduction libre de: However, we will only optimize the design and implementation of rubrics through scientific empirical research on benefits, as well as limitations.

En somme, la grille d'évaluation descriptive analytique a été étudiée à différents ordres d'enseignement, mais pas spécifiquement en adaptation scolaire au secondaire. C'est un outil qui permet de mettre en place une approche d'évaluation en soutien à l'apprentissage, mais qui peut être complexe à élaborer. De plus, selon une étude récente (Lamarche & Durand, 2022), les connaissances liées à l'utilisation de grilles d'évaluation descriptives analytiques sont minces chez les enseignants travaillant dans le contexte qui nous intéresse. Il s'avère ainsi intéressant d'approfondir le sujet.

L'objectif de notre article est de décrire la démarche de développement d'une grille d'évaluation descriptive analytique pour la discipline de Français, langue d'enseignement au premier cycle du secondaire, dans un contexte où des aménagements ont été apportés aux contenus pédagogiques pour tenter de répondre à la diversité des besoins et des profils des élèves présentant un TSA. Cette description permettra de déterminer les éléments fondamentaux nécessaires à la construction d'un tel outil dans un contexte si particulier.

Cadre de référence

L'interprétation critériée

Au Québec, selon la politique d'évaluation, l'interprétation critériée est à prioriser, de même que l'utilisation de la grille d'évaluation descriptive analytique (MEQ, 2003). Contrairement à l'interprétation normative, qui est beaucoup utilisée lors de la mesure des résultats d'apprentissage des élèves en vue de les comparer entre eux (Scallon, 2004), l'interprétation critériée permet de situer les manifestations de l'apprentissage des élèves en regard de standards (les attentes) axés sur le développement de la compétence (en progression, en cours d'apprentissage et de maîtrise, en fin d'apprentissage). Ce type d'interprétation favorise un jugement sur les productions et sur les performances des élèves à l'aide de critères pré-établis (Scallon, 2004; Simon & Forgette-Giroux, 2001). Avec un groupe d'élèves présentant un TSA, donc une grande diversité de profils cognitifs, l'interprétation critériée permet de se centrer sur ce qui est attendu pour chacun d'eux, car la comparaison entre les élèves est inutile, comme ils ne sont pas tous au même niveau scolaire.

Les constituants de la grille d'évaluation descriptive

La grille d'évaluation descriptive analytique est constituée de trois composantes : les critères, les descripteurs et les échelons.

Les critères d'évaluation prescrits dans les cadres d'évaluation (MELS, 2011) constituent les apprentissages qui seront évalués. Selon Durand et Trépanier (2011), les critères doivent être liés à la tâche, indépendants, peu nombreux et formulés de façon observable et mesurable. Leroux et Mastracci (2015) précisent que la rédaction de ces critères doit permettre d'observer la qualité de réalisation des élèves, et ce, de façon claire, grâce à un vocabulaire simple et compréhensible par le destinataire. Pour y parvenir, les critères sont reformulés sous forme d'indicateurs qui sont constitués des apprentissages à évaluer ainsi que de la dimension, de la qualité ou du point de vue observé.

Les critères sont ensuite répartis en différents niveaux de performance de la compétence, sur un continuum d'échelons. Selon Leroux et Mastracci (2015), il est préférable que le nombre d'échelons soit établi entre trois et six, en fonction de l'outil, de l'usage que l'on veut en faire, ainsi que de la complexité et de la nature de la tâche à évaluer. Les échelons peuvent suivre une notation alphanumérique (A, B, C et D ou 1, 2, 3 et 4) ou des expressions regroupées sur un même registre (p. ex., à sa façon de faire : très satisfaisant, satisfaisant, pas satisfaisant, insatisfaisant). Il s'agit ensuite de définir le critère de réussite, soit l'échelon attendu pour la réussite de la tâche et le seuil de réussite, c'est-à-dire ce qui est exigé minimalement en regard de la tâche.

Ce qui est attendu à chaque niveau de performance est développé à l'aide des descripteurs qui sont formulés de façon positive, exhaustive, juste et univoque. Ils décrivent ce que l'on peut observer dans la trace de l'élève, soit sa production ou sa performance. Ces descripteurs, en cohérence avec la performance attendue et le niveau donné, présentent une gradation régulière. Le vocabulaire à prioriser est aussi concis, précis et compréhensible. Une même structure rédactionnelle conforme aux aspects ciblés du critère assure une bonne modulation des manifestations observables en commençant par un verbe d'action, par exemple, *accorde correctement*, ou par l'objet évalué, par exemple, *les mots sont correctement...* Les erreurs les plus fréquemment observées dans la rédaction des descripteurs sont, premièrement, la mention de l'intervention de l'enseignant dans la performance de l'étudiant (p. ex., l'élève écrit une phrase complète avec l'aide

de son enseignant). Deuxièmement, la soustraction des éléments présents au niveau supérieur pour rédiger les niveaux inférieurs (p. ex., niveau A : l'élève écrit des phrases comprenant un sujet, un verbe et un complément ; niveau B : l'élève écrit des phrases comprenant un sujet et un verbe ; niveau C : l'élève écrit un sujet). Troisièmement, la rédaction d'un niveau insuffisant selon une liste d'erreurs typiques de la part de l'apprenant (p. ex., l'élève omet les marques d'accord, de pluriel et de ponctuation de base) (Bélaïr, 2014 ; Leroux, 2008, 2012 ; Tardif, 2004, 2006 cités dans Leroux et Mastracci, 2015).

Les étapes d'élaboration d'une grille d'évaluation

La construction d'une grille d'évaluation est un processus perfectible (Côté, 2014) qui se réalise de façon itérative en suivant quelques étapes. Stevens et Levi (2012) identifient quatre étapes clés :

- Réfléchir sur les attentes que nous avons envers les apprenants et ce que nous voulons qu'ils accomplissent ;
- Faire la liste des particularités de la tâche demandée et des objectifs d'apprentissage que nous voulons atteindre ;
- Regrouper nos attentes, rédiger nos critères d'évaluation à l'aide d'une dimension et identifier le nombre d'échelons nécessaires à notre grille ;
- Transférer nos critères dans la grille et rédiger les descripteurs qui seront placés dans la grille selon le nombre d'échelons.

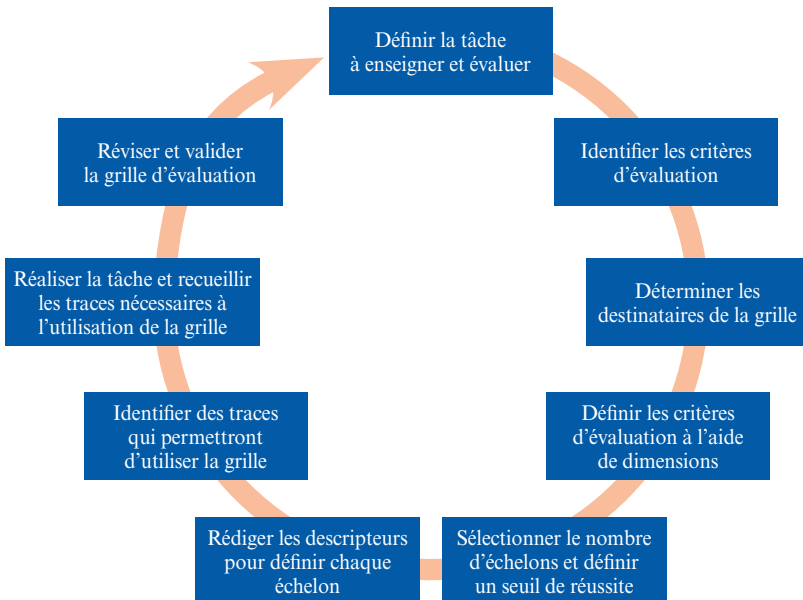
Quant à Durand et Chouinard (2012), ils présentent une séquence similaire, mais de façon beaucoup plus détaillée en proposant 12 étapes. Ces auteurs ajoutent à la sélection des critères d'évaluation l'identification du destinataire de la grille pour assurer un niveau de langage adapté, ce qui pourrait être fait à l'étape 1 de Stevens et Levi (2012). Ils séparent l'identification des indicateurs dans le critère d'évaluation et la sélection d'une dimension pour chacun des indicateurs en deux étapes distinctes, ce qui pourrait être fait à l'étape 3 de Stevens et Levi (2012). Ils ajoutent ensuite cinq étapes, lorsque la rédaction des descripteurs est terminée :

- Indiquer les traces des processus et des productions des élèves qui vont permettre la collecte de l'information ;
- Réviser la planification de l'évaluation de la situation d'apprentissage, si nécessaire ;

- Réaliser la situation d'apprentissage en informant les élèves des critères établis;
- Ramasser les copies des élèves et les interpréter à la lumière des descripteurs;
- Réviser la grille, si nécessaire. (Durand & Chouinard, 2012, p. 281)

À notre sens, certains éléments provenant des deux démarches à suivre sont pertinents. Alors que celle de Stevens et Levi (2012) semble très holistique et peu définie, celle de Durand et Chouinard (2012) semble trop précise et contraignante. Nous avons ainsi préféré établir une façon de faire qui reprend des éléments de chacun des auteurs et qui pourrait être utilisée dans différents milieux. La figure 1 présente le processus circulaire de la construction d'une grille.

Figure 1
Étapes de construction d'une grille d'évaluation descriptive



Selon Stiggins et al. (2006), une grille pertinente reprend adéquatement le contenu de la tâche, présente un vocabulaire compris de tous, est facile à utiliser, fiable et valide³. Durand et Mouffe (2014) ont repris ces éléments et les ont précisés à l'intérieur d'une métagrille permettant d'analyser la pertinence d'une grille descriptive analytique. Ainsi les auteures ont identifié six éléments centraux : 1) l'efficacité de l'organisation ; 2) la clarté du vocabulaire ; 3) la justesse des critères ; 4) la conformité des descripteurs ; 5) la pertinence des échelons et 6) l'adéquation du contenu. À cela, nous avons ajouté un septième élément spécifique à notre contexte, soit la qualité des aménagements à l'évaluation. Le tableau 1 présente la métagrille.

Méthode

Notre étude se situe dans une recherche plus vaste menée par les services régionaux de soutien et d'expertise (SRSE) à l'intention des élèves présentant un TSA ou des difficultés d'apprentissage de la région de l'île de Montréal. Ce projet a, entre autres, comme but de développer des ressources d'enseignement/apprentissage/évaluation spécifiques pour les élèves présentant un TSA et des défis scolaires importants. Ces ressources prennent la forme de situations d'apprentissage et d'évaluation (SAÉ). De ce fait, tout le matériel développé présente des contenus modifiés, soit moins exigeants que ce qui est attendu pour des élèves du secteur régulier du premier cycle du secondaire. Dans le présent article, nous nous concentrons sur le développement d'une grille d'évaluation descriptive analytique pour interpréter les traces d'apprentissage de ces élèves dans la matière Français, langue d'enseignement.

Nous avons choisi de suivre la démarche méthodologique de la recherche-développement de Harvey et Loiselle (2009). Nous avons pris une posture interprétative en utilisant les réflexions, les perceptions et les observations des participantes et de la chercheuse principale collectées lors de l'utilisation des prototypes de la grille. Ainsi, le but n'est pas de démontrer l'efficacité de notre instrument, mais bien d'identifier les caractéristiques essentielles que doit présenter une grille construite et utilisée dans le contexte d'un groupe ayant une diversité de besoins et de profils (Loiselle & Harvey, 2007). Suivant le modèle des auteurs, l'équipe du

3. Traduction libre de : Content (Does it cover everything?), Clarity (Does everyone understand what is meant?), Practicality (Is it easy to use by teachers and students?) and Technical quality/fairness (Is it reliable and valid?) (Stiggins et al., 2006, p. 203)

Tableau 1
Métagrille d'analyse

	A Exemplaire	B Pertinente	C À retravailler	D À refaire
1. Efficacité de l'organisation	La grille dresse un portrait global satisfaisant en donnant une vue d'ensemble des éléments recherchés. La grille est bien construite, condensée tout en étant complète, ce qui permet de remplir efficacement l'analyse de l'objet.	La grille est compréhensible, simple et les principaux éléments sont facilement repérables. La grille est bien construite condensée tout en étant complète, ce qui permet de remplir efficacement l'analyse de l'objet.	La grille est simple et ne donne qu'une partie des éléments recherchés. Ses limites nuisent à l'analyse de l'objet.	La grille est difficile à comprendre, peu d'éléments sont repérables et ses limites nuisent à l'analyse de l'objet.
2. Clarté du vocabulaire	Les énoncés sont formulés dans un langage univoque.	Les énoncés sont formulés dans un langage clair.	Les énoncés sont formulés dans un langage parfois ambigu.	Les énoncés sont formulés dans un langage imprécis, qui nuit à la compréhension.
3. Justesse des critères	Les critères précisent de façon compréhensible la dimension et l'objet évalué de manière mesurable ou observable.	Les critères identifient les dimensions, ils sont mesurables ou observables.	Les critères sont observables sans toutefois posséder de dimensions identifiées.	Les critères sont formulés de façon vague et imprécise, et les dimensions sont absentes.
	Les critères sont indépendants et ont tous une dimension appropriée.	Les critères sont indépendants et la même dimension est en progression d'un échelon à l'autre.	Les critères sont indépendants et la dimension peut varier d'un échelon à l'autre.	Les critères peuvent être dépendants les uns des autres et formulés uniquement sous forme quantitative.

	A Exemplaire	B Pertinente	C À retravailler	D À refaire
4. Conformité des descripteurs	Les descripteurs illustrent entièrement ce qui est visé. Ils sont formulés de façon positive et impartiale.	Les descripteurs illustrent partiellement ce qui est visé. Ils sont formulés de façon impartiale.	Les descripteurs illustrent difficilement ce qui est visé. Ils sont parfois formulés de façon négative.	Les descripteurs font abstraction de ce qui est visé. Ils sont formulés avec subjectivité.
5. Pertinence des échelons	La différence entre les échelons est évidente et tient compte d'une progression graduelle.	La différence entre les échelons est additive et montre un écart semblable de l'un à l'autre.	La différence entre les échelons est uniforme d'un critère à l'autre.	La différence entre les échelons est parfois absente et démontre une progression aléatoire.
6. Adéquation du contenu	Tous les descripteurs sont appropriés à la tâche à effectuer et permettent d'observer l'ensemble de la démarche d'apprentissage.	Les descripteurs sont appropriés à la tâche à effectuer, toutefois certains éléments pourraient être mieux formulés.	Les descripteurs sont appropriés à la tâche à effectuer, toutefois certains éléments sont superflus ou absents.	Le contenu de la tâche n'est pas approprié pour une évaluation selon une échelle descriptive : une seule bonne réponse est recherchée dans la plupart des éléments observés.
7. Qualité des aménagements à l'évaluation	Les aménagements sont identifiés séparément des apprentissages à réaliser. Ils sont spécifiques aux besoins de chacun des élèves et en lien avec le plan d'intervention.	Les aménagements sont identifiés et permettent l'accessibilité à la tâche selon les besoins des élèves.	Les aménagements sont génériques et inclus dans les performances des élèves.	L'identification des aménagements est floue ou absente.

Source : Grille adaptée de la métagrille de Durand et Mouffe (2014)

SRSE a réalisé les deux premières étapes, soit l'origine de la recherche et l'élaboration du cadre de référence. La problématique a été soulevée par des milieux scolaires et l'équipe a principalement utilisé des documents provenant du Massachusetts (Massachusetts Department of Elementary and Secondary Education, 2017) pour rédiger son référentiel, soit son cadre de référence et développer des SAÉ. Dans la présente section, nous abordons la méthodologie ainsi que le déroulement de l'opérationnalisation du projet. Nous présenterons les résultats par la suite. La figure 2 illustre le design du modèle de recherche.

Le déroulement de la recherche

L'étude s'est déroulée dans trois classes spécialisées de premier cycle du secondaire auprès d'élèves présentant un TSA provenant de centres de services scolaire (CSS) distincts, soit les trois CSS francophones de l'île de Montréal (CSS de Montréal, CSS de la Pointe-de-l'Île et CSS Marguerite-Bourgeoys). Dans chacune des classes, les enseignantes formées en adaptation scolaire ont été approchées par les conseillères pédagogiques (CP) du SRSE et ont volontairement accepté de participer au projet de recherche. Le tableau qui suit présente ces trois enseignantes.

Tableau 2
Présentation des participantes

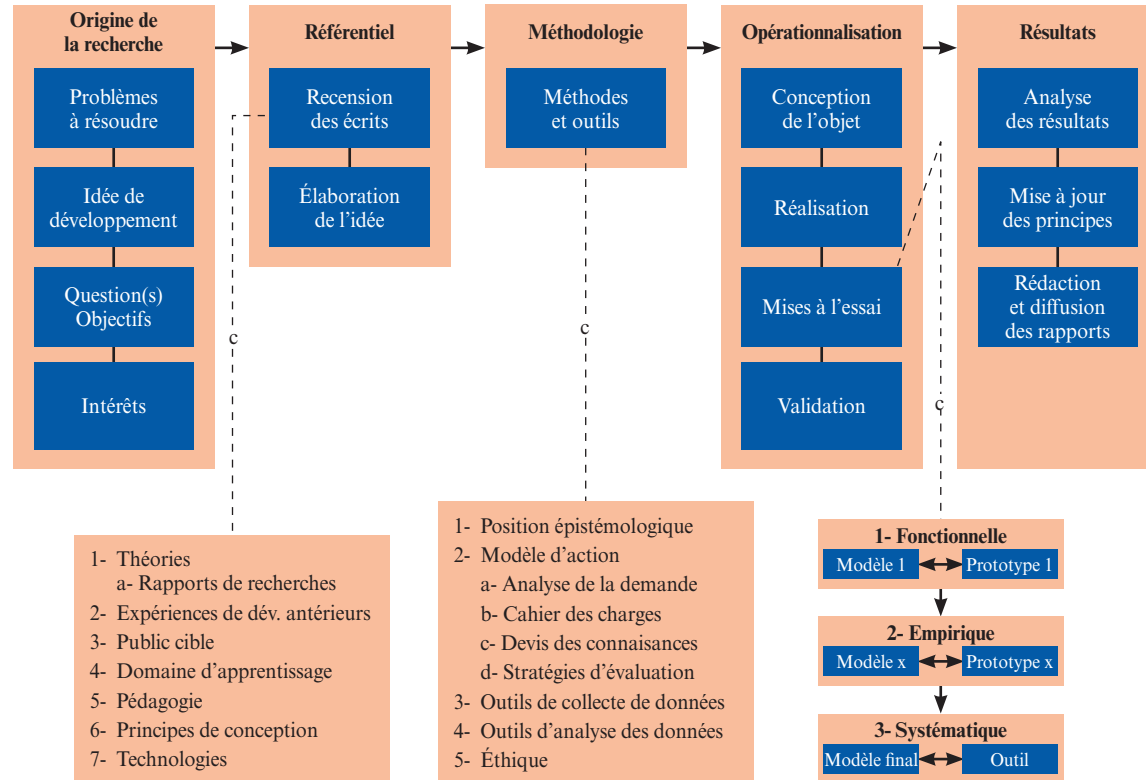
Participantes	Centre de services scolaire d'attache	Niveau d'enseignement	Nombre d'années d'expérience
Marie*	CSS Marguerite-Bourgeoys	1 ^{er} cycle du secondaire	4 ans
Raphaëlle	CSS de Montréal	1 ^{er} cycle du secondaire	5 ans
Alice	CSS de la Pointe-de-l'Île	1 ^{er} cycle du secondaire	2 ans

* Les prénoms utilisés sont des pseudonymes.

Pendant l'expérimentation, trois SAÉ⁴ comprenant chacune des grilles d'évaluation ont été mises à l'essai sur une période de trois années scolaires. Marie a participé au projet pendant deux ans. Elle était titulaire dans la même classe, mais les élèves ont changé d'une année à l'autre (classe 1a et 1b). Lors de la troisième année, elle a changé d'emploi pour occuper un poste de CP. Pendant les trois années, Raphaëlle a été titulaire dans la même classe et a vécu des changements d'élèves (classe 1a, 1b et 1c). Au

4. La SAÉ 1 portait sur le thème du chocolat, la SAÉ 2, sur le thème des boissons gazeuses et la SAÉ 3, sur le cinéma.

Figure 2
Modèle de recherche-développement en éducation



Source : Harvey et Loisel (2009, p. 110)

début de la troisième année, elle a mis un terme à sa participation pour des raisons de santé. Quant à Alice, elle était titulaire d'une classe lors de sa première année de participation (classe 1). Elle est devenue titulaire d'une classe complètement différente (classe 2) lors de sa deuxième année de participation. Elle est passée d'un groupe nécessitant des aménagements majeurs importants lors de l'apprentissage et de l'évaluation à un groupe où les élèves étaient identifiés pour réintégrer un groupe-classe ordinaire. Ils avaient donc droit à des aménagements mineurs. Sans être problématique, ce contexte nous a confrontées à une limite et à une force. Nous n'avons pas pu poursuivre la mise à l'essai de la grille dans une classe similaire d'une année à l'autre (classe de Raphaëlle). Or, lors de la troisième année, nous avons eu une approche plus collaborative avec Alice qui devait revoir complètement sa façon d'enseigner et d'évaluer ses élèves. Le tableau 3 présente les SAÉ réalisées par les enseignantes lors de leur participation au projet.

Tableau 3
Description des SAÉ réalisées lors des années de participation au projet

Participant	2018–2019	2019–2020	2020–2021
Marie	Classe 1a	Classe 1b	Deviens CP
	SAÉ 1	SAÉ 3	
	SAÉ 2	SAÉ 1	
Raphaëlle	Classe 1a	Classe 1b	Classe 1c
	SAÉ 1	SAÉ 3	Se retire
	SAÉ 2	SAÉ 2	du projet
Alice		Classe 1	Classe 2
		SAÉ 3	SAÉ maison
		SAÉ 1	

Le processus d'élaboration des grilles d'évaluation

En suivant le processus d'opérationnalisation de Harvey et Loiselle (2009), nous avons d'abord conçu et réalisé une grille, pour ensuite en faire la mise à l'essai et la validation (Harvey & Loiselle, 2009). Lors de la mise à l'essai, une attention particulière a été portée sur les caractéristiques pertinentes ou non du prototype utilisé. La validation de chacun des prototypes s'est faite auprès des participantes et d'un groupe d'experts (Brantlinger et al., 2005) à l'aide de fiche d'autoévaluation, d'entretiens semi-dirigés et de discussions. Le groupe d'experts était constitué de trois CP, d'une professeure universitaire et de la chercheuse principale. Ces

étapes ont été réalisées en continu sur trois prototypes de grilles. Nous avons commencé la conception des grilles à l'automne 2018, et le processus de développement s'est déroulé jusqu'à l'hiver 2021. Nous avons mis à l'essai trois grilles pendant les trois années scolaires. Le troisième prototype n'a pas pu être mis à l'essai à cause de la pandémie de COVID-19 au printemps 2020. En revanche, une forme aménagée de ce prototype a pu être mise à l'essai dans l'un des groupes. La figure 3 présente les étapes de développement de la grille.

Figure 3
Phases de la recherche-développement

Phase 1 2018-2019 Automne 2019	Phase 2 Hiver 2020	Phase 3 Printemps 2020* 2020-2021
<p>Conception/réalisation Prototype 1</p> <p>Mise à l'essai : SAÉ 1 Marie classe 1a Raphaëlle classe 1a SAÉ 2 Marie classe 1a Raphaëlle classe 1a</p> <p>SAÉ 3 (automne 2019) Marie classe 1b Raphaëlle classe 1b Alice classe 1</p> <p>Validation 1</p>	<p>Conception/réalisation Prototype 2</p> <p>Mise à l'essai : SAÉ 1 Marie classe 1b</p> <p>Poursuite mise à l'essai Prototype 1 SAÉ 1 Alice classe 1 SAÉ 2 Raphaëlle classe 1b</p> <p>Validation 2</p>	<p>Conception/réalisation Prototype 3</p> <p>Mise à l'essai : SAÉ maison Alice classe 2</p> <p>Validation 3 * Interruption du développement à cause de la COVID-19</p>

Les outils de collecte de données

Pour mener à bien le développement de notre instrumentation pour l'évaluation, nous avons eu recours à différents outils de collecte de données qualitatives (Loiselle & Harvey, 2007). Pendant les mises à l'essai de la grille, la chercheuse, qui était présente en classe, a inscrit ses observations dans un journal de bord. Les enseignantes ont rempli des fiches d'autoévaluation qui portaient sur leur utilisation de l'outil. Cette fiche comprenait des questions sur les différentes sections de la SAÉ et sur les pratiques des enseignantes, par exemple « J'ai été capable d'évaluer la compétence *Écrire des textes variés* auprès de mes élèves. ». Elles ont également participé à deux ou trois entretiens semi-dirigés qui abordaient l'utilisation de la grille d'évaluation. Ces entretiens comprenaient 13 questions, par exemple « Pouvez-vous me décrire les actes professionnels que vous adoptez à l'étape de l'interprétation dans la démarche d'évaluation? ». Les entretiens ont été systématiquement retranscrits pour en permettre l'analyse. Chaque étape de mise à l'essai a permis de consigner plusieurs informations quant à la pertinence la grille. C'est grâce à ces informations que des changements spécifiques ont été apportés à chacune des phases de validation.

La stratégie d'analyse de données

Les données recueillies ont été analysées au fur et à mesure afin d'assurer une compréhension complète de l'efficacité des prototypes utilisés. Une première analyse à l'aide de la métagrille a permis de cerner la pertinence de la grille d'évaluation. Une cote a été attribuée à chaque critère de la grille ce qui a permis d'émettre une appréciation globale du prototype. Nous avons également mené une analyse thématique de contenu (Van der Maren, 2004) pour le matériel qualitatif provenant des entretiens semi-dirigés et pour les fiches d'autoévaluation afin d'en faire ressortir les principaux thèmes pouvant nous guider sur les ajustements à apporter à l'instrument. Le tout a été appuyé par les observations consignées dans le journal de bord de la chercheuse principale. Ces analyses ont été réalisées à l'aide du logiciel *QDA Miner* (version 5.0.32) et ont été vérifiées et acceptées par le groupe d'experts participant à la validation des prototypes de la grille (Brantlinger et al., 2005). Ainsi, lorsque l'analyse a identifié un changement essentiel à apporter, l'équipe en a discuté au préalable. Ces discussions ont abordé les ajustements pour faire évoluer les prototypes. Chacune des phases de développement a été rigoureusement étudiée et a

permis de documenter en profondeur le développement de notre instrument pour l'évaluation de la discipline Français, langue d'enseignement auprès d'élèves du premier cycle du secondaire présentant un TSA.

Résultats

Les résultats obtenus sont présentés selon les phases de développement des grilles d'évaluation. Pour chaque phase, nous présentons une analyse du prototype appuyée d'exemples et de perceptions des participantes. Bien que le processus ait été réalisé pour les grilles des trois compétences de la discipline, dans le cadre du présent article, nous présentons seulement des extraits pour la compétence *Lire des textes variés*.

Phase 1

Les grilles d'évaluation ont été initialement élaborées pour accompagner des activités d'apprentissage spécifiques pour les élèves présentant un TSA. Ces activités ont été conçues en intégrant des aménagements majeurs apportés au contenu d'apprentissage. Pour y parvenir, l'équipe de recherche a dû redéfinir les compétences scolaires en fonction du PFEQ (MELS, 2006), développer des activités pédagogiques incluant un guide pour l'enseignant ainsi que tout le matériel nécessaire pour la réalisation des activités par les élèves. L'instrumentation pour l'évaluation comprenait des grilles d'évaluation descriptives analytiques pour les différentes compétences ciblées en français ainsi que des activités d'autoévaluation et de coévaluation par les pairs entre élèves.

La prise en compte d'exigences moins élevées, soit les aménagements majeurs au contenu pour les compétences, ont permis d'offrir des attentes plus accessibles aux élèves. Ainsi, une déclinaison des éléments de compétences pour la discipline Français, langue d'enseignement diminuait le niveau de complication, tout en conservant le niveau de complexité des activités d'apprentissage. Cette déclinaison, nommée « points d'entrée », a servi à la conception de l'instrumentation pour l'évaluation en concordance avec les activités d'apprentissage. Les points d'entrée ont été rédigés en amont de tout le matériel pédagogique, afin d'avoir un continuum développemental pour les compétences scolaires, avec des niveaux de développement plus simples permettant d'atteindre la zone proximale de développement de certains élèves. L'équipe de recherche s'est inspirée des travaux du département de l'éducation primaire et secondaire

du Massachusetts (2017) pour concevoir un continuum aménagé et en cohérence avec le PFEQ (MELS, 2006). C'est avec ces attentes plus accessibles que les tâches d'apprentissage et l'instrumentation à l'évaluation ont été conçues. Plusieurs versions des grilles d'évaluation ont été réalisées, validées et mises à l'essai pendant la première phase de conception. Les premières versions ont utilisé un vocabulaire provenant davantage des points d'entrée et du continuum aménagé. Pour illustrer ces premières tentatives, nous présentons un extrait de la grille pour la compétence *Lire des textes variés* en exemple, dans le tableau 4.

Les premières versions de grilles étaient chargées et comprenaient des éléments qui alourdissaient l'outil, dont des informations superflues. Les colonnes « critère d'évaluation modifié » et « manifestation du critère » et les titres « composante de la compétence », « points d'entrée » et « habiletés d'accès » se sont avérés inutiles. Afin de rendre l'instrument plus accessible, nous avons consulté les participantes pour mieux cerner les besoins de leurs élèves et leur niveau de développement de la compétence. Les observations en classe consignées dans le journal de bord de la chercheuse principale ont également permis d'ajuster la grille et de proposer des niveaux plus représentatifs de la réalité vécue en classe. Le tableau 5 présente un extrait de la version améliorée de la grille au terme de la conception/réalisation de la phase 1.

Ce premier prototype a été mis à l'essai par deux enseignantes au printemps 2019 et ensuite par les trois enseignantes à l'automne 2019. Raphaëlle et Alice ont bien reconnu le niveau de leurs élèves dans ce prototype de grille et ont eu de la facilité à l'utiliser, ce qui a permis de brosser un portrait juste du niveau de compétence des élèves tout en étant conforme avec le PFEQ (MELS, 2006) et en répondant à un besoin mentionné par les enseignantes :

La grille d'évaluation pour vraiment avoir une image globale, et je trouve que c'est un très beau *wrap up* et que je ne suis plus obligée d'aller chercher dans mes trucs [outils didactiques et documentation ministérielle]. (Alice, entretien 1)

Justement, le lien avec le PFEQ que moi je ne fais pas vraiment, là, ça le fait quand même, ça épure la compétence, ça garde de quoi de précis, et c'est comme très décortiqué après. À différents niveaux, donc oui, c'est plus mon lien en français, c'est plus parlant pour le lien avec le PFEQ. (Raphaëlle, entretien 1)

Tableau 4

Extrait d'une grille d'évaluation pour la compétence Lire des textes variés en cours de réalisation

SAÉ 1: Chocolat que c'est bon!							
Compétence: Lire des textes variés							
Critère d'évaluation	Critère d'évaluation modifié	Composante de la compétence visée (Intermédiaire)	Manifestation du critère	Points d'entrée			Habilités d'accès
				A	B	C	
Critère 2 Interprétation fondée	L'élève formule, généralement dans ses mots, une interprétation à partir des textes lus et commet quelques maladresses.	L'élève fait des liens entre ses connaissances antérieures et nouvelles par rapport à la fabrication du chocolat.	Contenu	L'élève fait part de ses connaissances nouvelles par rapport à la fabrication du chocolat.	L'élève identifie des mots nouveaux dans le texte en lien avec la fabrication du chocolat.	L'élève identifie des images qui sont nouvelles pour lui (qui ne viennent pas du texte) qui représentent la fabrication du chocolat.	L'élève identifie une image en lien avec la fabrication du chocolat lorsqu'elle est présentée avec une image hors contexte.

Tableau 5
Extrait du prototype 1 pour l'évaluation de la compétence Lire des textes variés

SAÉ 3: Critique de cinéma							
Compétence : Lire des textes variés							
Critère d'évaluation	Critère d'évaluation de la SAÉ	A L'élève...	B L'élève...	C L'élève...	D L'élève...	E L'élève...	F L'élève...
Critère 2 Interprétation fondée	Qualité des explications Activité 5	apporte des explications claires, pertinentes et précises quant à son appréciation du film.	apporte des explications de façon plus ou moins claire.	apporte quelques explications à la demande de l'enseignante.	émet un commentaire global à la demande de l'enseignante.	pointe le regroupement « j'aime » et émet un commentaire global à la demande de l'enseignante.	pointe parmi les éléments qu'il a aimés, celui qu'il a préféré en l'associant à un motif positif (illustration).

Par ailleurs, Marie avait des élèves avec des profils scolaires très hétérogènes dont certains n'avaient pas accès à des aménagements majeurs, mais plutôt à des aménagements mineurs. Dans ce cas, les descripteurs ne répondaient pas aux attentes pour certains élèves. L'échelle ne présentait pas un continuum suffisamment large pour situer tous les portraits des élèves de cette classe, ce qui a causé des problèmes d'interprétation chez l'enseignante. De plus, pour elle, l'utilisation même de la grille était ardue. Ce n'était pas un outil que Marie avait l'habitude d'utiliser et elle éprouvait des difficultés à bien cerner son utilité :

Quand je lis les grilles, tu sais, je trouve tellement que c'est mince entre chacun [des descripteurs]. Là, il faut que j'aie voir le mot qui ne revient pas dans l'énoncé pour être sûre. (Marie, entretien 1)

Dans ce passage, Marie relève un obstacle important en lien avec la clarté du vocabulaire, soit le critère 2 de la métagrille. Malgré l'attention particulière portée aux mots dans la rédaction des descripteurs, la complexité de la compétence et la diversité des profils des élèves rendent l'exercice de rédaction compliqué. Il devient irréaliste d'atteindre une compréhension univoque de l'instrument. Toutefois, en utilisant un vocabulaire simple et des mots discriminants, nous pouvons aspirer à une compréhension de la majorité.

Selon les critères de la métagrille, ce prototype semble pertinent pour deux enseignantes et «à retravailler» pour la troisième. Ceci explique pourquoi Raphaëlle et Alice ont continué d'utiliser ce premier prototype à l'hiver 2020. Pour Marie, nous avons convenu d'ajouter des échelons à la grille afin de répondre à une plus grande variété d'apprenants. Nous avons également revu les descripteurs et tenté de les rendre plus distinctifs. Le tableau 6 présente une synthèse des changements apportés.

Les aménagements mis en place pendant la tâche n'ont pas été abordés dans cette phase de développement. Les enseignantes étaient plutôt concentrées sur les compétences qu'elles devaient enseigner et, surtout, sur leur appropriation du matériel pédagogique. Ce n'est qu'à la phase 3 que nous avons entrepris la réflexion concernant l'introduction de ces aménagements.

Tableau 6
Résultats de l'analyse de la mise à l'essai du prototype 1

Participants	Appréciation globale	Aspects à retravailler
Raphaëlle et Alice	B – Pertinent, mais il n'a pas été question des aménagements apportés.	C7 – Documenter les aménagements apportés à l'évaluation.
Marie	D – À refaire	C2 – Clarifier le vocabulaire. C3 – Le choix limité des critères ne permet pas d'avoir une vision de toute la compétence. C4 – Les descripteurs ne permettent pas d'évaluer un élève en aménagement mineur. C5 – Les échelons ne sont pas assez nombreux pour être représentatifs de tous les élèves. C7 – Documenter les aménagements apportés à l'évaluation.

Phase 2

Le deuxième prototype présente une grille à laquelle des échelons ont été ajoutés, afin de couvrir la diversité de développement des apprenants, à l'aide de descripteurs qui répondent davantage aux attentes du PFEQ. Toutefois, les tâches d'apprentissage étant développées pour des élèves présentant des défis cognitifs élevés, il a été difficile de rédiger ces nouveaux descripteurs, car les productions attendues restaient simples. De ce fait, nous avons tenté de concevoir une grille qui se rapprochait de l'instrumentation utilisée dans les classes ordinaires, ce qui représentait un défi supplémentaire. Le tableau 7 présente un extrait de cette grille.

L'intention de ce deuxième prototype était d'inclure tous les portraits d'élèves d'un groupe alors que certains ont droit à des aménagements majeurs et d'autres non. Rapidement, nous avons constaté que nous ne pouvions pas avoir des attentes de niveau secondaire 1 avec des tâches simples. Alors que nous tentions de rendre plus complexes les attentes, les descripteurs s'éloignaient de plus en plus des apprentissages travaillés en classe. De plus, l'ajout d'échelons a rendu la grille plus lourde, ce qui a nui à l'analyse du développement de la compétence. Le tableau 8 présente les éléments à retravailler de ce deuxième prototype.

Tableau 7
Extrait du prototype 2 pour l'évaluation de la compétence Lire des textes variés

SAÉ 1: Chocolat que c'est bon!									
Compétence : Lire des textes variés									
Critère d'évaluation	Critère d'évaluation de la SAÉ	3 L'élève...	2 L'élève...	1 L'élève...	0 L'élève...	00 L'élève...	A L'élève...	B L'élève...	C L'élève...
Critère 2 Interprétation fondée	Qualité des liens établis	fait de nombreux liens judicieux entre les éléments importants clés sur la fabrication du chocolat et les connaissances acquises.	fait des liens pertinents entre plusieurs éléments clés de la fabrication du chocolat et les connaissances acquises.	fait des liens entre les éléments généraux sur la fabrication du chocolat et les connaissances acquises.	fait des liens simples entre les éléments importants de la fabrication du chocolat et les connaissances acquises.	fait des liens généralement corrects entre les éléments principaux de la fabrication du chocolat.	fait des liens entre les connaissances antérieures et nouvelles au sujet de la fabrication du chocolat.	fait des liens, avec l'aide de l'enseignante, entre ses connaissances antérieures et nouvelles au sujet de la fabrication du chocolat.	fait part de ses nouvelles connaissances au sujet de la fabrication du chocolat.

Tableau 8
Résultats de l'analyse de la mise à l'essai du prototype 2

Participant	Appréciation globale	Aspects à retravailler
Marie	C – À retravailler	<p>C1 – La grille est chargée, comprend trop d'éléments, notamment un très grand nombre d'échelons.</p> <p>C4 – Les descripteurs illustrent difficilement ce qui est travaillé en classe.</p> <p>C6 – Les descripteurs sont appropriés à la tâche réalisée, toutefois la tâche n'est pas appropriée pour tous les élèves.</p> <p>C7 – Les aménagements sont inclus dans les descripteurs.</p>

Dans cette grille, comme dans celle du prototype 1, nous avons inclus l'aide apportée par l'enseignante dans les descripteurs. Par exemple, dans l'échelon B, nous trouvons : « fait des liens, avec l'aide de son enseignante, entre ses connaissances antérieures et nouvelles [...] ». Nous avons beaucoup discuté de cette façon de faire avec les participantes, car elles constataient que certains élèves pouvaient démontrer un niveau de compétence plus faible, mais de façon autonome (p. ex., l'élève fait des liens entre ses connaissances antérieures et nouvelles), alors que d'autres manifestaient des niveaux de compétence plus élevés, mais avec de l'aide (p. ex., l'élève fait des liens pertinents entre plusieurs éléments clés). Nos descripteurs étaient trop hermétiques et cela créait de la confusion dans le jugement des enseignantes, notamment au moment de la notation.

Oui, il va être en modifié parce que je suis toujours là. Tu sais, ça, c'est clair pour moi. Si je suis là, si j'enlève des questions... Tu sais, je connais un peu les limites, mais le situer entre 60 et 100, parce que c'est ça nos bulletins, je le mets où? (Marie, entretien 1)

Au terme de cette deuxième phase de validation, nous avons remarqué que l'inclusion d'un grand nombre de descripteurs dans une seule grille était une tâche inefficace. Lorsque le niveau des attentes varie beaucoup d'un élève à l'autre, il devient essentiel de présenter des tâches différentes, selon les besoins des élèves. Cela implique également des instruments différents pour en tenir compte. Dans le cadre de notre projet, les SAÉ proposées comportaient des aménagements majeurs, notamment sur le plan du contenu et de la complexité. Il était ainsi difficile d'observer un

niveau de première secondaire pour une compétence. De plus, le fait que l'aide apportée par l'enseignant soit incluse dans les descripteurs a créé un obstacle dans l'utilisation de l'instrument.

Phase 3

Nous avons conçu notre troisième prototype au printemps 2020, lors du confinement occasionné par la pandémie de COVID-19. L'équipe a pris le temps d'analyser les différentes données collectées depuis le début de l'expérimentation pour en arriver aux constats suivants : il était préférable de revenir à un nombre d'échelons moindre, spécifiquement rédigé pour des élèves dont les attentes étaient modifiées, et de distinguer le niveau d'autonomie de l'élève de son niveau d'appropriation du contenu. En effet, le fait de séparer l'aide apportée et les contenus scolaires permettait de mieux documenter le type d'aménagement mis en place. Étaient-ce des aménagements liés à la production (forme, contenu, niveau de complexité, etc.) ou au soutien (aide technologique, aide de l'enseignante, etc.)? Le tableau 9 présente un extrait du troisième prototype.

Ce prototype n'a pas pu être mis à l'essai dans les mêmes contextes que les prototypes précédents. Parmi les trois participantes, seule Alice était en poste pendant l'année 2020-2021, et son groupe-classe ne nécessitait que des aménagements mineurs, contrairement aux élèves initialement concernés par notre matériel. Compte tenu de nos observations de l'année précédente, nous avons conseillé à l'enseignante de ne pas utiliser le matériel développé par le SRSE. Cela dit, avec la participation d'Alice, nous avons pris le temps de coconstruire une nouvelle grille d'évaluation en fonction de ses besoins et suivant les principes acquis pendant l'expérimentation. La première grille a été élaborée pour évaluer la compétence *Écrire des textes variés* de niveau premier cycle du secondaire et comprenant tous les critères d'évaluation. Nous avons rédigé des descripteurs spécifiques à la tâche et nous avons ajouté un espace pour que l'enseignante indique les aménagements mineurs utilisés pendant l'activité. Le tableau 10 présente un extrait de cette grille qui a plu à l'enseignante.

Inclure les différentes adaptations [aménagements] ou des commentaires ciblés faisant état de l'aide ou du soutien réel donné à l'élève lors de la tâche permet une meilleure compréhension commune de son niveau d'autonomie par rapport à une tâche scolaire de son niveau scolaire. (Alice, fiche réflexive, SAÉ maison)

Tableau 10

Extrait du prototype 3 aménagé au contexte d’Alice pour évaluer la compétence Écrire des textes variés

		A	B	C	D	E
		L’élève...	L’élève...	L’élève...	L’élève...	L’élève...
Adaptation à la situation d’écriture	Respect du sujet 30	décrit plusieurs aspects d’un personnage dont son physique et sa personnalité à l’aide d’idées très bien développées.	décrit les aspects physiques et psychologiques d’un personnage à l’aide d’idées bien développées.	décrit sommairement les aspects physiques et psychologiques d’un personnage, les idées sont imprécises.	décrit les aspects physiques et psychologiques brièvement, ne fournit aucune précision.	décrit des aspects décousus ou reprend des idées sans les reformuler dans ses mots.
		30	24	18	12	6
<i>Aménagements mineurs apportés</i>						

La troisième phase de validation des grilles a été interrompue par la pandémie de COVID-19. Nous n'avons pas pu consolider la mise à l'essai de ce prototype avec le matériel pédagogique du SRSE, mais les informations obtenues par la troisième enseignante étaient encourageantes.

La discussion qui suit présente les bons coups et les éléments à retravailler énoncés à travers notre processus de développement. De plus, nous proposerons une analyse globale de l'utilisation des grilles d'évaluation par les participantes et des lignes directrices pour la création de grilles d'évaluation descriptives analytiques pour les élèves du secteur de l'adaptation scolaire.

Discussion

L'objectif de notre étude était de développer une grille d'évaluation spécifique pour les compétences en Français, langue d'enseignement (lecture, écriture, communication orale) pour les élèves du premier cycle du secondaire présentant un TSA.

Les premiers constats concernent l'hétérogénéité des profils des apprenants, et ce, malgré le diagnostic unique de TSA (Poirier et al., 2017). Au sein d'un même groupe d'élèves, les profils sont variés et les besoins sont différents, ce qui appuie la nécessité d'avoir une interprétation critériée de l'évolution des compétences scolaires pour ces élèves, soit de comparer les productions à des critères et non les unes en comparaison des autres (Simon & Forgette-Giroux, 2001). Toutefois, la création d'un instrument d'évaluation spécifique reste un processus ardu. Ce constat présente une limite importante dans la flexibilité de l'instrument identifié par Goodrich (1996). Contrairement à ce que soulève cette auteure, une trop grande hétérogénéité entre les apprenants devient un obstacle important à la conception d'une grille d'évaluation descriptive. Pour bien cerner les difficultés rencontrées, l'utilisation systématique de la métagrille a été favorable. Cet instrument nous a permis d'identifier rapidement et de synthétiser les éléments à retravailler lors de l'évaluation de la pertinence de nos prototypes.

Tout d'abord, nous avons retravaillé la pertinence et le nombre d'échelons, tout comme l'adéquation des contenus dans les deux premiers prototypes. Comme le mentionnent Leroux et Mastracci (2015), un nombre trop élevé d'échelons ne permet pas une évaluation efficace du niveau de compétence. Il est préférable de maintenir un nombre limité d'échelons

et d'ajuster la tâche demandée aux élèves pour que ces derniers puissent bien démontrer leur niveau de compétence. Pour ce faire, il est nécessaire d'établir des dispositifs de différenciation qui respectent le rythme de développement des élèves, en mettant en place des aménagements majeurs lors de l'apprentissage et de l'évaluation.

Ensuite, la clarté du vocabulaire a été un élément à retravailler. Effectivement, Marie a clairement évoqué le fait que le vocabulaire utilisé dans les grilles pouvait être difficile à comprendre ; il était donc essentiel de choisir un vocabulaire simple et compréhensible pour la majorité des destinataires (p. ex., les parents et les élèves eux-mêmes). À ce sujet, Wiertz et al. (2020) soulignent qu'une construction conjointe de l'outil permet cette compréhension, ce que nous avons également observé lors de l'élaboration de la dernière grille. En effet, Alice y ayant participé, il était beaucoup plus facile pour elle de l'utiliser et de bien saisir les descripteurs.

Un élément important de notre grille est l'identification distincte des aménagements mis en place. En effet, comme l'avaient identifié Leroux et Mastracci (2015), il est préférable de ne pas inclure les aménagements apportés dans les descripteurs d'une grille. Or, le fait de laisser un espace libre pour une rédaction plus personnalisée des aménagements serait aussi un élément favorable à ajouter. Alice semble d'ailleurs avoir apprécié cette façon de faire. Pour d'autres, il s'agirait d'un endroit pour laisser des commentaires.

Au terme de cette analyse, nous dégagons trois principes essentiels à considérer lors de la création de grilles d'évaluation descriptives analytiques pour des compétences disciplinaires des élèves présentant un TSA :

- Maintenir un nombre raisonnable de descripteurs pour une tâche donnée. Pour ce faire, la création de sous-groupes au sein même de la classe permettrait d'évaluer le niveau de compétence des élèves se situant sur un continuum similaire ;
- Utiliser un vocabulaire simple et compréhensible. Pour y parvenir, la coconstruction avec des enseignants semble être une approche favorable ;

- Distinguer les aménagements apportés de la description des manifestations observables. Dans un contexte d'aménagements majeurs, les descripteurs sont déjà des traces des aménagements proposés, mais il est également intéressant de documenter le niveau de soutien apporté aux élèves. Dans un contexte d'aménagements mineurs, il semble intéressant de laisser un espace réservé à l'identification des aménagements mis en place ou pour des commentaires.

Conclusion

Alors que notre étude présente un regard novateur sur la grille d'évaluation descriptive analytique, il faut relever certaines limites. Tout d'abord, l'interruption du processus de développement n'a pas permis d'établir une dernière analyse complète de l'outil. De plus, les principales mises à l'essai ont été conduites au cours d'une même année scolaire et auprès d'un petit échantillon. Il serait donc intéressant de reprendre les principes identifiés dans notre étude et de les expérimenter dans un nouveau contexte qui inclut des élèves présentant un TSA, en proposant une mise à l'essai sur une période assez longue, avec un soutien extérieur fourni par un spécialiste ou par un CP. Également, il aurait été pertinent d'inclure les enseignantes dans l'élaboration de la grille dès le début du processus. Considérant que ce fut un succès avec Alice en fin de projet, cela aurait pu être bénéfique pour l'entièreté de notre démarche.

La grille d'évaluation descriptive analytique constitue un instrument intéressant à utiliser pour juger du niveau de compétence des élèves présentant un TSA. Or, à plusieurs reprises, une participante a soulevé le point suivant :

Tu sais, tu me donnerais deux semaines, je partirais avec mes huit élèves et je te dresserais des portraits de feu. J'utiliserais la progression des apprentissages pour les situer et identifier ce qu'ils ont réussi. Je ferais juste ça et j'aurais zéro problème à dire s'il est en secondaire 1 ou secondaire 1 modifié. Mais, je n'ai pas ce temps-là. (Marie, entretien 1)

Cette enseignante indique vouloir faire une interprétation encore plus descriptive que celle que permet la grille, soit une évaluation narrative (Margrain, 2010). Les autres participantes ont également soulevé ce point et se disent plus à l'aise pour rédiger des commentaires pour communiquer le niveau de compétence de leurs élèves. De ce fait, un outil comme le portfolio pourrait bonifier l'efficacité de la grille d'évaluation (Derycke,

2000). Enseigner à des élèves présentant un TSA est un contexte pour lequel les pratiques évaluatives sont encore peu développées (Lamarche & Durand, 2021), ce qui est aussi le cas pour l'ensemble des types d'élèves du secteur de l'adaptation scolaire (Branciard et al., 2016). Alors que le système scolaire québécois valorise l'inclusion des élèves en classe ordinaire, il est essentiel d'approfondir les connaissances concernant les pratiques évaluatives auprès des élèves qui nécessitent des aménagements mineurs ou majeurs afin d'assurer une évaluation juste, égale et équitable pour tous.

Révision linguistique : Marie-Claire Legaré

Mise en page : Emmanuel Gagnon

Résumé en portugais : Eusébio André Machado

Réception : 20 décembre 2022

Version finale : 19 juillet 2023

Acceptation : 26 août 2023

LISTE DES RÉFÉRENCES

- Baribeau, A. (2009). Analyse des pratiques d'évaluation des compétences d'enseignants de français, secondaire premier cycle, pour établir un jugement professionnel lors du bilan des apprentissages [mémoire de maîtrise, Université du Québec à Trois-Rivières]. Cognition. <https://depot-e.uqtr.ca/id/eprint/1187>
- Baribeau, A. (2015). Analyse des pratiques d'évaluation d'enseignants du secondaire IV et V dans des décisions sommatives de certification des apprentissages des élèves [thèse de doctorat, Université du Québec à Trois-Rivières, Université du Québec à Montréal]. Cognition. <https://depot-e.uqtr.ca/id/eprint/8161>
- Berthiaume, D., David, J., & David, T. (2011). Réduire la subjectivité lors de l'évaluation des apprentissages à l'aide d'une grille critériée : repères théoriques et applications à un enseignement interdisciplinaire. *Revue internationale de pédagogie de l'enseignement supérieur*, 27(2), 1-12. <https://doi.org/10.4000/ripes.524>
- Branciard, L., Mias, C., & Benoit, H. (2016). Vers une évaluation scolaire inclusive? *La nouvelle revue de l'adaptation et de la scolarisation*, 74(2), 58. <https://www.cairn.info/revue-la-nouvelle-revue-de-l-adaptation-et-de-la-scolarisation-2016-2-page-5.htm>
- Brantlinger, E., Jimenez, R., Klingner, J., Pugach, M., & Richardson, V. (2005). Qualitative Studies in Special Education. *Exceptional Children*, 71(2), 195-207. <https://doi.org/10.1177/001440290507100205>
- Brind'Amour, A. (2018). Les pratiques évaluatives d'enseignantes de 1^{re} année du primaire à l'égard de la compétence en lecture : portrait d'une commission scolaire québécoise [mémoire de maîtrise, Université du Québec à Rimouski]. Sémaphore. <https://semaphore.uqar.ca/id/eprint/1473/>
- Brookhart, S. M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education*, 3. <https://www.frontiersin.org/articles/10.3389/educ.2018.00022>
- Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, 67(3), 343368. <https://doi.org/10.1080/00131911.2014.929565>
- Calhoun, M. B., Fuchs, L. S., & Hamlett, C. L. (2000). Effects of Computer-Based Test Accommodations on Mathematics Performance Assessments for Secondary Students with Learning Disabilities. *Learning Disability Quarterly*, 23(4), 271-282. <https://doi.org/10.2307/1511349>
- Chan, Z., & Ho, S. (2019). Good and bad practices in rubrics: the perspectives of students and educators. *Assessment & Evaluation in Higher Education*, 44(4), 533545. <https://doi.org/10.1080/02602938.2018.1522528>
- Chaumont, M. (2015). Documenter le jugement professionnel d'enseignantes et d'enseignants lors de l'évaluation certificative de compétences qui se développent dans plus d'un cours au collégial. [essai de maîtrise, Université de Sherbrooke]. Savoirs UdeS. <http://hdl.handle.net/11143/8966>
- Côté, F. (2014). Construire des grilles d'évaluation descriptives au collégial : guide d'élaboration et exemples de grille. PUQ.
- Dawson, P. (2017). Assessment rubrics: towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, 42(3), 347-360. <https://doi.org/10.1080/02602938.2015.1111294>

- Deaudelin, C., Desjardins, J., Dezutter, O., Thomas, L., Corriveau, A., Lavoie, J., Bousadra, F., & Hébert, M. (2007). L'évaluation formative en contexte de renouveau pédagogique au primaire : analyse de pratiques au service de la réussite. *Nouveaux cahiers de la recherche en éducation*, 10(1), 27-45. <https://doi.org/10.7202/1016856ar>
- Denis, C., Lison, C., & Lépine, M. (2016). Pratiques d'adaptation de l'évaluation pour l'inclusion d'élèves dyslexiques : qu'arrive-t-il au deuxième cycle du secondaire? *La nouvelle revue de l'adaptation et de la scolarisation*, (74), 27-39. <https://doi.org/10.3917/nras.074.0027>
- Derycke, M. (2000). La grille critériée et le portfolio à l'épreuve du suivi pédagogique. *Revue française de pédagogie*, 132(1), 23-32. <https://doi.org/10.3406/rfp.2000.1030>
- Di Lalla, J. (2017). Le rapport au jugement professionnel dans l'évaluation du savoir-écrire du texte argumentatif en français, langue maternelle : des cas d'enseignants de 5^e secondaire [essai, Université de Sherbrooke]. Savoirs UdeS. <https://savoirs.usherbrooke.ca/handle/11143/11896>
- Dolz-Mestre, J., & Tobola Couchepin, C. (2015). Les pratiques d'évaluation dans l'enseignement de la production écrite d'un genre argumentatif à l'école primaire. *Le Français dans le monde. Recherches et applications*, (58), 144-153. <https://archive-ouverte.unige.ch/unige:93376>
- Durand, M.-J., & Chouinard, R. (2012). L'évaluation des apprentissages : de la planification de la démarche à la communication des résultats (2^e éd.). Marcel Didier.
- Durand, M.-J., & Mouffe, M. (2014). La métagrille, ou outil pour évaluer la qualité des grilles d'évaluation : exemple d'une production en histoire. Dans M.-J. Durand et N. Loye (dir.), *L'instrumentation pour l'évaluation. La boîte à outils de l'enseignant évaluateur* (p. 6992). Marcel Didier.
- Durand, M.-J., & Trépanier, I. (2011). Validité des situations de compétence : élaboration d'une grille d'analyse. Dans G. Raïche, K. Paquette-Coté et D. Magis (dir.), *Des mécanismes pour assurer la validité de l'interprétation de la mesure en éducation. Volume 2, L'évaluation*. Presses de l'Université du Québec.
- Frith, U. (2003). *Autism: Explaining the enigma*. Blackwell Publishing.
- Fuchs, L. S., & Fuchs, D. (2001). Helping Teachers Formulate Sound Test Accommodation Decisions for Students with Learning Disabilities. *Learning Disabilities Research & Practice*, 16(3), 174-181. <https://doi.org/10.1111/0938-8982.00018>
- Gérard, F.-M. (2010). L'évaluation des compétences à travers des situations complexes. Dans G. Baillat, D. Niclot et D. Ulma (dir.), *La formation des enseignants en Europe : approche comparative* (p. 231-241). De Boeck Supérieur.
- Goodrich, H. (1996). Understanding Rubrics. *Educational Leadership*, 54(4). <https://www.ascd.org/el/articles/understanding-rubrics>
- Harvey, S., & Loiselle, J. (2009). Proposition d'un modèle de recherche développement. *Recherches qualitatives*, 28(2), 95-117. <https://doi.org/10.7202/1085274ar>
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Jung, L. A., & Guskey, T. R. (2011). *Grading Exceptional and Struggling Learners*. Corwin.
- Lamarque, M.-A., & Durand, M.-J. (2021). Synthèse des écrits scientifiques portant sur les pratiques d'évaluation des apprentissages actuelles auprès d'élèves du secondaire ayant un trouble du spectre de l'autisme. *Revista Educativa - Revista de Educação*, 23(1), 8601. <https://doi.org/10.18224/educ.v23i1.8601>

- Lamarche, M.-A., & Durand, M.-J. (2022). Développement du sentiment d'efficacité personnelle perçu par des enseignantes œuvrant auprès d'élèves ayant un trouble du spectre de l'autisme au secondaire et appropriation de nouvelles pratiques évaluatives. *Canadian Journal of Education/Revue canadienne de l'éducation*, 45(3), 835866. <https://doi.org/10.53967/cje-rce.v45i3.5239>
- Leroux, J. L. (2010). *Analyse des pratiques évaluatives d'enseignantes et d'enseignants dans une approche par compétences au collégial* [thèse de doctorat, Université de Sherbrooke]. Savoirs UdeS. <http://savoirs.usherbrooke.ca/handle/11143/947>
- Leroux, J. L., & Mastracci, A. (2015). *Une méta grille d'évaluation à expérimenter*. Association québécoise de pédagogie collégiale (AQPC). <https://eduq.info/xmlui/handle/11515/35822>
- Loiselle, J., & Harvey, S. (2007). La recherche développement en éducation : fondements, apports et limites. *Recherches qualitatives*, 27(1), 40–59. [http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero27\(1\)/loiselle.pdf](http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero27(1)/loiselle.pdf)
- Louis, R. (2004). L'évaluation des apprentissages, théorie et pratique (2^e éd.). Beauchemin.
- Margrain, V. (2010). Narratives of Young Gifted Children. *Kairaranga*, 11(2), 33–38. <https://eric.ed.gov/?id=EJ925414>
- Martin, E., Lefrançois, C., Guichard, A., Tapp, D., & Arsenaault, L. (2016). Processus de coconstruction d'une grille critériée pour l'évaluation de productions écrites complexes à l'université. *Revue internationale de pédagogie de l'enseignement supérieur*, 32(2), 1-22. <https://doi.org/10.4000/ripes.1094>
- Massachusetts Department of Elementary and Secondary Education. (2017). Resource Guide to the 2017 Massachusetts Curriculum Frameworks for Students with Disabilities: English language arts and literacy. Massachusetts Department of Elementary and Secondary Education
- Mastracci, A. (2011). Des critères d'évaluation génériques et une grille d'évaluation à échelles descriptives globales pour évaluer des apprentissages en créativité au collégial [essai de maîtrise, Université de Sherbrooke]. Savoirs UdeS. <http://hdl.handle.net/11143/9682>
- Meirieu, P. (2016). L'école, mode d'emploi : des « méthodes actives » à la pédagogie différenciée (16^e éd.). ESF éditeur.
- Ministère de l'Éducation, du Loisir et du Sport (MELS). (2006). *Programme de formation de l'école québécoise Enseignement secondaire, premier cycle*. Gouvernement du Québec. http://www.education.gouv.qc.ca/fileadmin/site_web/documents/PFEQ/prfrmseclercycle2.pdf
- Ministère de l'Éducation, du Loisir et du Sport (MELS). (2011). *Cadre d'évaluation des apprentissages : Français, langue d'enseignement, enseignement secondaire 1^{er} et 2^e cycle*. Gouvernement du Québec. http://www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/formation_jeunes/francais-langue-d-enseignement-sec.pdf
- Ministère de l'Éducation, du Loisir et du Sport (MELS). (2014). Document d'information. Précisions sur la flexibilité pédagogique, les mesures d'adaptation et les modifications pour les élèves ayant des besoins particuliers. Gouvernement du Québec. <https://www.cssdgs.gouv.qc.ca/wp-content/uploads/2016/10/Pr%C3%A9cisions-sur-la-flexibilit%C3%A9-p%C3%A9dagogique-les-mesures-dadaptation-et-les-modifications-pour-les-%C3%A9l%C3%A8ves-ayant-des-besoins-particuliers-MEES.pdf>
- Ministère de l'Éducation du Québec (MEQ). (2003). *Politique d'évaluation des apprentissages*. Gouvernement du Québec. http://www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/evaluation/13-4602.pdf

- Ministère de l'Éducation du Québec (MEQ). (2001). *Programme de formation de l'école québécoise. Éducation préscolaire et enseignement primaire*. Gouvernement du Québec. <http://www.education.gouv.qc.ca/enseignants/pfeq/primaire/>
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 1291-144. <https://doi.org/10.1016/j.edurev.2013.01.002>
- Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, 30, 1-19. <https://doi.org/10.1016/j.edurev.2020.100329>
- Paré, M. (2012). *Pratiques d'individualisation en enseignement primaire au Québec visant à faciliter l'intégration des élèves handicapés ou des élèves en difficulté au programme de formation générale* [thèse de doctorat, Université de Montréal]. Papyrus. <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/6293>
- Poirier, N., Abouzeid, N., Taieb-Lachance, C., & Smith, E.-L. (2017). Le programme de formation et les stratégies éducatives déclarées offerts aux adolescents ayant un trouble du spectre de l'autisme qui fréquentent une classe spécialisée. *Canadian Journal of Education*, 40(4), 457-485. <https://journals.sfu.ca/cje/index.php/cje-rce/article/view/2524>
- Ramoo, L. D., & Durand, M.-J. (2016). Documenter les façons de faire d'enseignants de 6^e année du primaire en mathématiques, en lecture et en écriture dans toutes les étapes de la démarche d'évaluation. *Canadian Journal of Education/Revue canadienne de l'éducation*, 39(4), 124.
- Salend, S. J., & Duhaney, L. M. G. (2002). Grading students in inclusive settings. *Teaching Exceptional Children*, 34(3), 8-15.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Édition du Renouveau Pédagogique Inc.
- Scanlon, D., & Baker, D. (2012). An Accommodations Model for the Secondary Inclusive Classroom. *Learning Disability Quarterly*, 35(4), 212-224. <https://doi.org/10.1177/0731948712451261>
- Simon, M., & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. *Practical Assessment, Research, and Evaluation*, 7(1), 1-4. <https://doi.org/10.7275/bh4d-me80>
- Stevens, D. D., & Levi, A. J. (2012). *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning* (2^e éd.). Stylus Publishing, LLC.
- Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2006). *Classroom assessment for student learning. Doing it right - using it well* (2^e éd.). Educational Testing Services.
- Tardif, J. (1993). L'évaluation dans le paradigme constructiviste. Dans R. Hivron (dir.), *L'évaluation des apprentissages: réflexions, nouvelles tendances et formation* (p. 27-56). Éditions du CRP.
- Van der Maren, J.-M. (2004). *Méthodes de recherche pour l'éducation* (2^e éd.). De Boeck & Larcier.
- Vermeulen, P. (2014). *Comment pense une personne autiste?* Dunod. <https://doi.org/10.3917/dunod.verme.2014.01>

- Vincent, K. (2017). Conception de grilles d'évaluation pour l'évaluation formative et certificative des compétences de stagiaires de Soins infirmiers au collégial [essai de maîtrise, Université de Sherbrooke]. Savoirs UdeS. <http://hdl.handle.net/11143/14055>
- Wiertz, C., Van Mosnenck, S., Galand, B., & Colognesi, S. (2020). Évaluer l'oral quand on est enseignant ou chercheur : points de discussion et prises de décision dans la coconception d'une grille critériée. *Mesure et évaluation en éducation*, 43(3), 1-37. <https://doi.org/10.7202/1083006ar>
- Yetiş, V. A. (2017). Les grilles d'évaluation critériée pour évaluer des performances: exemples pour la production écrite. *Journal of Uludag University Faculty of Education*, 30(2), 683-703. <http://hdl.handle.net/11452/12963>