Mesure et évaluation en éducation

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

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

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 and Micheline-Joanne Durand

Volume 45, Number spécial, 2022

Translation Issue

Submitted: December 20, 2022
Final version: July 19, 2023
Accepted: August 26, 2023

URI: https://id.erudit.org/iderudit/1110994ar
DOI: https://doi.org/10.7202/1110994ar

See table of contents

Publisher(s)

ADMEE-Canada

ISSN

0823-3993 (print)
2368-2000 (digital)

Explore this journal

Cite this article

Lamarche, M.-A. & Durand, M.-J. (2022). 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. *Mesure et évaluation en éducation*, *45*(spécial), 121–156. https://doi.org/10.7202/1110994ar

Article abstract

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.

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[1]

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

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

1. The French version was published in issue 45(3) 2022: https://doi.org/10.7202/1107235ar

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.

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.

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.

Note des auteurs : La correspondance liée à cet article peut être adressée à marie-aimee.lamarche@umontreal.ca.

Research problem

Since the advent of the Quebec Education Program or QEP (Ministère de l'Éducation du Québec (MEQ), 2001), a competency-based approach has been implemented in the province of Quebec, with a focus on "integrating the (theoretical and practical) knowledge, skills and attitudes necessary to successfully complete complex tasks that are meaningful to the student and necessary for subsequent adjustment to adult life" (Louis, 2004, p. 22, translated freely). In this context, complex tasks are used not merely to measure the knowledge a student has learned, but to be able to observe the construction of knowledge, i.e., manifestations of cognitive and metacognitive strategies (Gérard, 2010; Tardif, 1993). Interpreting "proof" of students' learning in this type of task requires specific instrumentation that makes use of properly formulated and agreed-upon criteria. "Analytic and developmental evaluation grids" (Berthiaume et al., 2011), commonly referred to in the English-language literature as "rubrics" (Stevens & Levi, 2012), are defined as a tool supporting the evaluator's judgment on the quality and progression of a performance in accordance with specific criteria and descriptors that can be associated with a weighting or rating notation of some kind (Brookhart, 2018; Brookhart & Chen, 2015; Panadero & Jonsson, 2020; Yetis, 2017). Given their transparency and precision, rubrics help ensure a better understanding of the different manifestations of student learning by placing them on a continuum.

Descriptive rubrics are of growing interest to researchers and give rise to a diversity of subjects of study. Some authors focus on how to understand and use rubrics (Baribeau, 2009, 2015;Chan & Ho, 2019; Dawson, 2017; Wiertz et al., 2020), others on what makes rubrics an effective tool (Brookhart & Chen, 2015; Durand & Trépanier, 2011). In this article, we focus on rubric development from the perspective of assessment that supports learning in the context of special education and students living with specific cognitive conditions, such as autism spectrum disorder (ASD). This disorder manifests, among other things, through atypical functioning in certain circumstances. For example, these students may have difficulty showing empathy toward others (see theory of mind); they may

find it difficult to process information in its entirety without dwelling on details (see central coherence), and they may experience difficulty getting organized, so that they require routine activities (see executive functions) (Frith, 2003; Vermeulen, 2014). To our knowledge, the construction of descriptive rubrics in such a context has never been documented in the scholarly research (Lamarche & Durand, 2021).

Assessment of students with an autism spectrum disorder with autism

In the Quebec school system, assessment frameworks involve evaluation criteria that are not always suited to the diverse profiles of students with ASD. These students have atypical learning and cognitive profiles and do not perceive life in the same way as neurotypical young people (Vermeulen, 2014). As a result, it becomes difficult to use the same evaluation criteria to judge their learning. In this context, teachers can use instructional accommodations that may or may not influence their judgment. For example, they can use digital tools that support the students' mental processes and do not influence their judgment, such as speech recognition. They can also simplify or remove evaluation criteria that have a direct influence on evaluative judgment. This distinction is explored further in the next section. According to Jung and Guskey (2011), the student accommodations set out in an individualized education plan (IEP) help guide the design of assessment instruments, i.e., rubrics. Furthermore, the authors point out that interpreting the productions of students entitled to accommodations represents an additional challenge for teachers, who must distinguish between the student's learning and their degree of autonomy in carrying out the task. A rubric appears to be an appropriate tool for supporting learning and incorporating the accommodations required for the academic progress of students with ASD.

Task accommodations

Many students with autism exhibit significant cognitive challenges that necessitate major adjustments in teaching practices, both in terms of learning and assessment, and especially the use of differentiated instructional strategies.

English-speaking authors (Calhoon et al., 2000; Fuchs & Fuchs, 2001; Salend & Duhaney, 2002; Scanlon & Baker, 2012) refer to these mechanisms of differentiation using the terms "inclusive setting," "adaptations" or "test accommodations." In the French-language academic literature, authors (Denis et al., 2016; Meirieu, 2016; Paré, 2012; Poirier et al., 2017)

use terms such as *accommodations*, *adaptations*, *aménagements* or *pratiques différenciées*. All of these terms entail adapting practices with a view to integrating students into a "regular" context. In addition, Quebec ministerial documents (2014) identify three levels of differentiated intervention: pedagogical flexibility (common interventions that are developed according to the diversity of students), adaptive measures (designed to attenuate barriers to learning without lowering requirements) and modification measures (interventions that require an IEP and are designed to change the expectations and requirements under the QEP (Ministére de l'Education, du Loisir et du Sport (MELS), 2006)).

In our project, teachers had to differentiate their assessment methods according to the learning content, and modify their expectations. We use the term "major accommodations" when the content taught and assessed falls below the requirements of the education program specific to the age group. In contrast, when differentiation mechanisms help attenuate obstacles to learning without modifying the level required, we use the term "minor accommodations."

Rubrics are difficult to use with students with ASD because the learning that takes place in class does not meet the evaluation criteria prescribed by the ministry (Lamarche & Durand, 2022; MELS, 2011). Little research has investigated the development of rubrics for academic competencies for this type of learner. Although various teaching and assessment practices have been documented (Poirier et al., 2017), the interpretation of competency assessment remains a significant challenge (Lamarche & Durand, 2021, 2022).

Literature review

Rubrics have been the subject of various studies that focus on how they are used as well as their advantages and challenges. While some authors directly address the use of this tool at the college level (Chaumont, 2015; Côté, 2014; Leroux, 2010; Mastracci, 2011; Vincent, 2017) and university level (Berthiaume et al., 2011; Martin et al., 2016), the research conducted at the elementary and secondary school levels addresses the tool through teachers' assessment practices. Ramoo and Durand (2016) documented the assessment practices of 6th grade teachers. Their findings indicated that teachers were more likely to use uniform numerical and descriptive-scale rubrics during writing activities than during mathematics and reading activities. In her study, Brind'Amour (2018) documented the assessment

practices of 1st grade elementary teachers for the competency "Reads a variety of texts." Although the descriptive rubric is one of the instruments that can be used, the results show that the teachers instead used an observation grid. Deaudelin et al. (2007) documented the formative assessment practices of 13 elementary teachers. Their results showed that the teachers favoured the use of custom rubrics to support self-assessment, group work and problem-solving in mathematics. This instrument is used in order to get students involved in their own assessment and, above all, in their own learning. At the secondary school level, Di Lalla (2017) and Baribeau (2009, 2015) observed teachers' practices at the end of an education cycle. In both contexts, teachers used the rubrics provided by the government during ministry exams to document their judgment. These studies show that rubrics are used in Quebec classrooms, although none of the studies examined the development of the rubrics, nor did they address the special education sector where instructional accommodations are necessary, even though this instrument could be appropriate in this context.

Goodrich (1996) argues that a rubric is a flexible instrument that helps take into account the heterogeneous profiles of students in the same class. Indeed, depending on students' abilities, expectations can be adjusted by adding or subtracting levels of performance from the rubric (Goodrich, 1996). In addition, literature reviews have clearly shown that rubrics support learning and the development of self-regulation in students and teachers, and can positively influence student motivation (Brookhart & Chen, 2015; Jonsson & Svingby, 2007; Panadero & Jonsson, 2013). Among other things, these positive elements result from transparency, as documented by Berthiaume et al. (2011) in a university context within a history course. Goodrich points out that a rubric enables students to better understand the task at hand, while also enabling teachers to better plan their teaching.

For their part, Martin et al. (2016) and Wiertz et al. (2020) highlight how time-consuming constructing a descriptive rubric can be (the instrument does not save time), as well as the importance of having a clear understanding of its content. According to these authors, a rubric is a complex instrument that is not always easy to use, given the lexicon it contains. Wiertz et al. (2020) note that, to overcome this obstacle, it helps to build rubrics with others as a team and to share ideas. This was also found by Dolz-Mestre and Tobola Couchepin (2015), who included students in developing a rubric.

In their recent literature review, Panadero and Jonsson (2020) explore criticism of rubrics in general. The results of their review identify six themes[2] which, according to these authors, show that the empirical evidence on which these criticisms are based is unsound and largely rooted in anecdotal events. They also point to a lack of understanding of the concept by the authors of the various studies making up the corpus they examined. They conclude that rubrics can be used and conceptualized in many different ways. Nevertheless, according to these authors, "we will only optimize the design and implementation of rubrics through scientific empirical research on benefits, as well as limitations" (Panadero & Jonsson, 2020, p. 17).

In short, rubrics have been studied at different education levels, but not specifically in special education at the secondary school level. They can be used to implement an assessment approach that supports learning, but they can be complex to develop. Moreover, according to a recent study (Lamarche & Durand, 2022), teachers working in this context have limited knowledge of how to use these rubrics. This would be a subject worth exploring in greater depth.

The aim of this article is to describe the process of developing a rubric for the subject of French (language of instruction) in middle school, in a context where adaptations have been made to instructional content to respond to the diverse needs and profiles of autistic students. This description will help determine the fundamental elements required to build such a tool in this particular context.

Reference framework

Criterion-based interpretation

According to Quebec's assessment policy, criterion-based interpretation should be preferred, as should be the use of rubrics (MEQ, 2003). Contrary to the normative interpretation, which is often used when measuring students' learning results with a view to comparing them with one another (Scallon, 2004), criterion-referenced interpretation makes it possible to situate manifestations of student learning in relation to standards

2. "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 & Jonsson, 2020, p. 14).

(expectations) associated with competency development (beginning, learning and mastery, and completion of learning). This type of interpretation favours a judgment on productions and performances using pre-established criteria (Scallon, 2004; Simon & Forgette-Giroux, 2001). With a group of students with ASD who presented with a wide variety of cognitive profiles, criterion-referenced interpretation helps focus on what is expected for each profile, as between-student comparisons are useless given that they are not all at the same educational level.

Components of a rubric

A rubric has three components: criteria, descriptors and levels of performance.

The evaluation criteria prescribed in assessment frameworks (MELS, 2011) constitute the learning that will be evaluated. According to Durand and Trépanier (2011), criteria must be task-specific, independent, few in number and worded in a way that makes them observable and measurable. Leroux and Mastracci (2015) clarify that these criteria must be written so as to enable the evaluator to observe the quality of the students' performance, and to do so clearly, using simple vocabulary that the recipient will be able to understand. To this end, the criteria are reformulated into indicators, which are made up of the learning to be assessed, as well as the dimension, the quality or the point of view observed.

The criteria are then divided into different levels along a continuum of levels of performance. According to Leroux and Mastracci (2015), it is preferable to set the number of levels of performance between three and six, depending on the tool, its intended use, and the nature and complexity of the task to be assessed. The levels of performance can use alphanumeric notation (A, B, C and D or 1, 2, 3 and 4) or expressions grouped together in the same register (e.g., way of doing things: exceptional, satisfactory, underdeveloped, limited). The next step is to define the criterion for success, i.e., the level of performance expected for the task and the passing threshold, i.e., what is minimally required in relation to the task.

What is expected at each level of performance is developed using descriptors, which are worded in a positive, exhaustive, fair and unambiguous way. They describe what can be observed in the "proof" of the student's learning i.e., their production or performance. These descriptors, consistent with the expected performance and given level, present a regular gradation. The vocabulary should be concise, precise and

comprehensible. Using one consistent writing structure in line with the targeted aspects of the criterion will help properly modulate observable manifestations, for example, starting with an action verb, e.g., *conjugates correctly*, or with the object of assessment, for example, *the words are correctly*.... The most frequently observed errors in the writing of descriptors are, first, mentioning the teacher's intervention in the student's performance (e.g., the student writes a complete sentence with the teacher's help). The second is subtracting elements present in the higher level when writing the lower levels (e.g., level A: student writes sentences containing a subject, a verb and a complement; level B: student writes sentences containing a subject and a verb; level C: student writes a subject). The third is writing an unsatisfactory level, based on a list of typical learner errors (e.g., student omits markers of agreement, plurals and basic punctuation) (Bélair, 2014; Leroux, 2008, 2012; Tardif, 2004, 2006 cited in Leroux and Mastracci, 2015).

Steps in developing a rubric

Constructing a rubric is a perfectible process (Côté, 2014) that is carried out iteratively, in several stages. Stevens and Levi (2012) identify four key steps:

- Reflecting on the expectations we have of learners and what we want them to achieve;
- Listing the specifics of the task and the learning objectives to be achieved;
- Grouping together expectations, drafting the evaluation criteria according to a dimension and identifying the number of levels of performance required for the rubric; and
- Transferring the criteria to the rubric and writing the descriptors that will be entered into the rubric according to the number of levels of performance.

For their part, Durand and Chouinard (2012) present a similar sequence, but in much greater detail, involving 12 steps. In addition to selecting evaluation criteria, these authors also cite identifying the rubric's recipient to ensure an appropriate language level, which could be done in step 1 in accordance with Stevens and Levi (2012). They separate the identification of indicators in the evaluation criterion and the selection of a dimension for each indicator into two separate steps, which could be done in step 3 of Stevens and Levi (2012). They add five steps once the descriptors have been drafted:

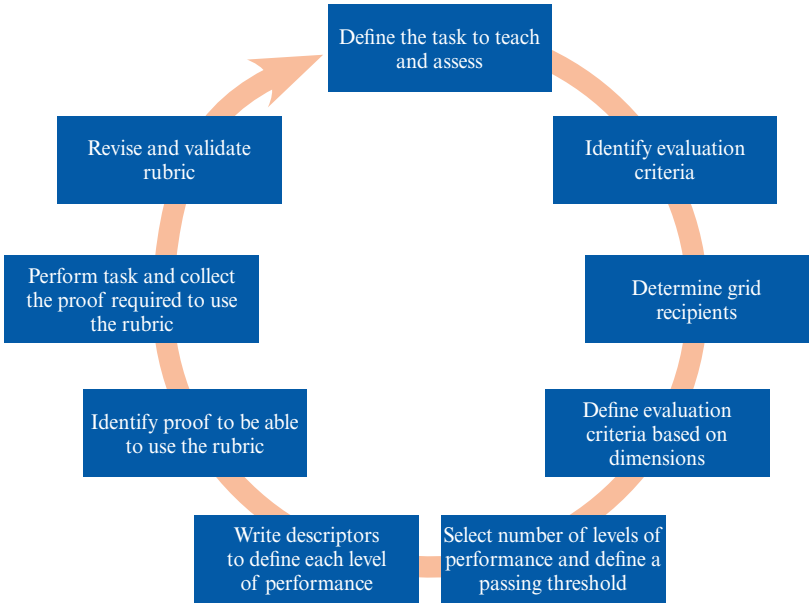
- Indicating proof of the students' processes and productions that will enable information gathering;

– Reviewing the assessment plan for the learning situation, if necessary;
– Conducting the learning situation while informing the students of the criteria established;
– Collecting student copies and interpreting them in light of the descriptors; and
– Revising the rubric if necessary (Durand & Chouinard, 2012, p. 281).

In our view, certain elements of both approaches are relevant. While Stevens and Levi's (2012) approach seems very holistic and loosely defined, Durand and Chouinard's (2012) approach appears to be overly precise and restrictive. We therefore preferred to establish an approach that incorporates elements from each of these authors and that could be used in different settings. Figure 1 shows the circular process of rubric construction.

According to Stiggins et al. (2006), a suitable rubric adequately captures the contents of the task, exhibits easily understandable vocabulary, and is

Figure 1
Steps in developing a descriptive rubric



Source: Figure inspired by Durand and Chouinard (2012) and Stevens and Levi (2012)

easy to use, reliable and valid[3]. Durand and Mouffe (2014) have taken up these elements and refined them within a metarubric for analyzing the relevance of a rubric. These authors identified six central elements: 1) effective organization; 2) clear vocabulary; 3) appropriate criteria; 4) relevant descriptors; 5) suitable levels of performance and 6) appropriate content. To these we added a seventh element specific to our context, namely the quality of assessment-related accommodations. This metarubric is set out in Table 1.

Methodology

Our study is part of a larger research project conducted by the Services régionaux de soutien et d'expertise (SRSE) for students with ASD and learning disabilities in the Montreal area. One of the aims of this project is to develop specific teaching/learning/assessment resources for students with ASD and significant academic challenges. These resources take the form of learning and evaluation situations (LES). As a result, all the material developed contains modified content, i.e., it is less demanding than what is expected of students in regular middle school. In this article, we focus on the development of a rubric for interpreting these students' proof of learning in the subject of French (language of Instruction).

We have chosen to follow Harvey and Loiselle's (2009) methodological approach of research and development. We adopted an interpretative stance, using the reflections, perceptions and observations of the participants and the principal researcher collected during the implementation of the rubric prototypes. Thus, the aim is not to demonstrate the effectiveness of our instrument, but rather to identify the essential characteristics that must be included in a rubric constructed and used in the context of a group with a diversity of needs and profiles (Loiselle & Harvey, 2007). In keeping with the authors' model, the SRSE team carried out the first two phases, i.e., the origin of the research and the development of the reference framework. The question of the difficulty in evaluating the competencies of students with ASD was raised by teachers themselves. The team mainly used documents from Massachusetts (Massachusetts Department of Elementary and Secondary Education, 2017) to draft its reference framework and develop the learning and evaluating situations (LESs). In this section, we discuss the methodology and how the project was implemented.

We then present the results. Figure 2 illustrates the research model design.

3. "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)

Table 1
Metarubric

	A **Exemplary**	**B** **Satisfactory**	**C** **Needs work**	**D** **Needs to be done over**
1. Effective organization	Rubric paints a satisfactory overall picture, providing a comprehensive overview of the desired elements. Rubric is well-constructed, condensed yet complete, making it easy to effectively analyze the object.	Rubric is simple and easy to understand, and the main elements are easily identifiable. The rubric is well-constructed, condensed yet complete, making it easy to analyze the object.	Rubric is simple and provides only a portion of the information required. Its limitations are detrimental to analyzing the object.	The rubric is difficult to understand, there are few identifiable elements, and its limitations get in the way of analyzing the object.
2. Clear vocabulary	Statements are worded in unambiguous language.	Statements are worded in clear language.	Statements are worded in sometimes ambiguous language.	Statements are worded in imprecise language, which hinders comprehension.
3. Sound criteria	The criteria clearly specify the dimension and the object being evaluated in a measurable or observable way.	The criteria identify dimensions, and are measurable or observable.	The criteria are observable but lack identified dimensions.	The criteria are formulated in a vague and imprecise way, and the dimensions are missing.
	The criteria are independent and all have an appropriate dimension.	The criteria are independent, and the same dimension progresses from one level to the next.	The criteria are independent and the dimension can vary from one level to another.	Criteria can be mutually dependent and worded in quantitative form only.
4. Relevant descriptors	The descriptors fully illustrate what is intended. They are worded in a positive and impartial way.	The descriptors partially illustrate what is intended. They are worded impartially.	The descriptors hardly illustrate what is intended. They are sometimes worded in a negative way.	Descriptors neglect what is intended. They are worded subjectively.

	A Exemplary	B Satisfactory	C Needs work	D Needs to be done over
5. Appropriate levels of performance	The difference between levels of performance is clear and reflects a gradual progression.	The difference between levels is additive and shows a similar gap from one to the other.	The difference between levels is uniform from one criterion to the next.	The difference between levels is sometimes absent, demonstrating a haphazard progression.
6. Match of criteria	All descriptors are appropriate to the task at hand and make it possible to observe the entire learning process.	The descriptors are appropriate for the task at hand, but some elements could be better worded.	The descriptors are appropriate for the task at hand, but some elements are superfluous or missing.	The content of the task is not suitable for assessment on a descriptive scale: a single correct answer is sought in most of the elements observed.
7. Assessment accommodations	Accommodations are identified separately from the learning to be achieved. They are specific to the needs of each student and linked to the individualized education plan.	Accommodations are identified and make the task accessible according to the students' needs.	Accommodations are generic and included in student performance.	Identification of accommodations is unclear or absent.

Source: Rubric adapted from Durand and Mouff (2014)

Research process

The study took place in three middle school special education classes for students with ASD belonging to three different school service centers (CSS) on the island of Montreal (CSS de Montréal, CSS de la Pointe-de-l'Île and CSS Marguerite-Bourgeoys). In each classroom, teachers trained in special education were approached by educational advisors and voluntarily agreed to take part in the research project. These three teachers are presented in the table below.

Table 2
Presentation of the participants

Participants	School services centre	Level of instruction	No. of years of experience
Marie*	CSS Marguerite-Bourgeoys	Middle school	4 years
Raphaëlle	CSS de Montréal	Middle school	5 years
Alice	CSS de la Pointe-de-l'Île	Middle school	2 years

* The first names used here are pseudonyms.

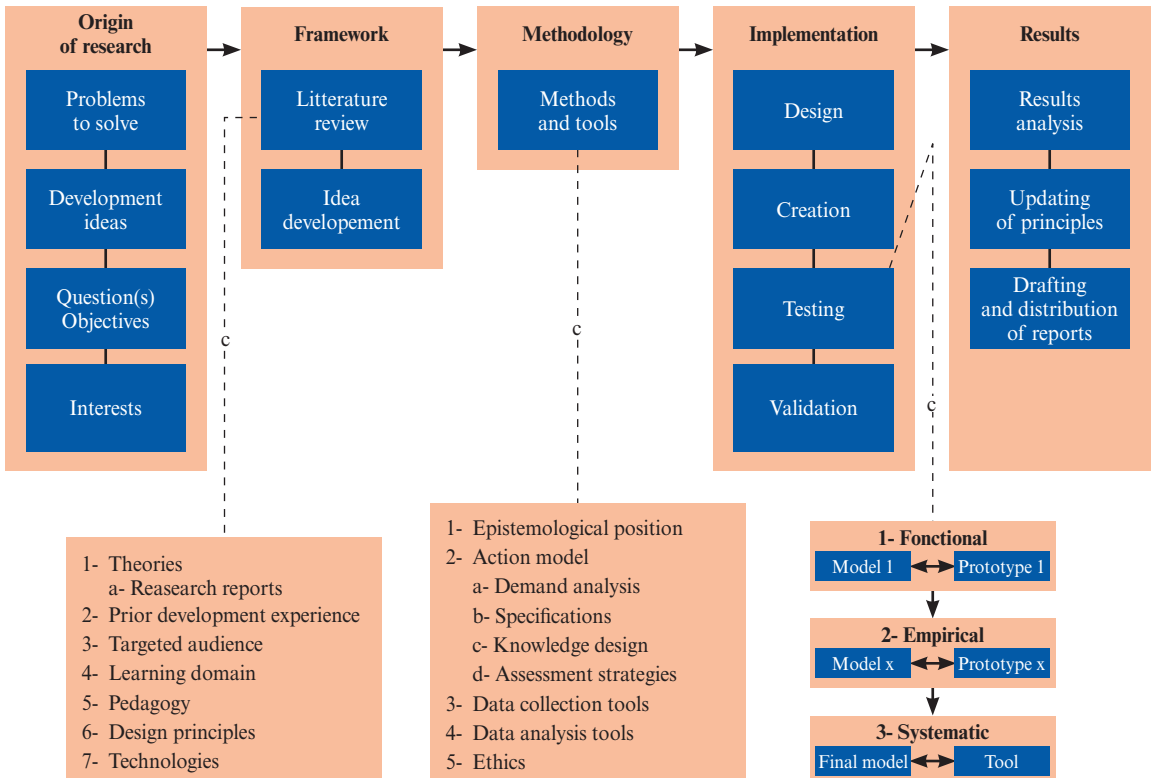
During experimentation, three LESs,[4] each with its own rubric, were tested over a period of three academic years. Marie took part in the project for two years. She was in charge of the same class, but the students changed from one year to the next (class 1a and 1b). In the third year, she changed jobs to take up a position as an educational advisor. For all three years, Raphaëlle was the teacher in charge of the same class, with some changes in her students (class 1a, 1b and 1c).

At the beginning of the third year, she withdrew for health reasons. For her part, Alice was a classroom teacher in her first year of participation (class 1). She then became responsible for a completely different class (class 2) in her second year of participation. She went from a group requiring major accommodations in terms of learning and assessment to a group where students were identified for reintegration into a regular class group. They were therefore entitled to minor accommodations. Although not problematic, this context did confront us with a limitation and a strength. We were unable to continue testing the rubric in a similar class from one year to the next (Raphaëlle's class). In the third year, however,

4. LES 1 was on the topic of chocolate, LES 2, on the topic of soft drinks and LES 3, on cinema.

Figure 2

Model of development research in education



Source: Harvey and Loiselle (2009, p. 110)

we took a more collaborative approach with Alice, who had to completely rethink her way of teaching and assessing her students. Table 3 shows the LESs carried out by the teachers during their participation in the project.

Table 3
Description of the LESs carried out during years of participation in the project

Participants	2018–2019	2019–2020	2020–2021
Marie	Class 1a LESs 1 LESs 2	Class 1b LESs 3 LESs 1	Became ed. advisor
Raphaëlle	Class 1a LESs 1 LESs 2	Class 1b LESs 3 LESs 2	Class 1c Withdrew from project
Alice		Class 1 LESs 3 LESs 1	Class 2 Custom LESs

The rubric development process

Consistent with the operationalization process suggested by Harvey and Loiselle (2009), we first designed and built a rubric for testing and validation (Harvey & Loiselle, 2009). During testing, special attention was paid to relevant and irrelevant characteristics of the prototype used. The prototypes were validated with the participants as well as a panel of experts (Brantlinger et al., 2005), using self-evaluation forms, interviews and discussions. The panel consisted of three educational advisors, a university professor and the principal investigator. All three began designing the rubrics in fall 2018, and the development process ran until winter 2021. We tested three rubrics over the three school years. The third prototype could not be tested due to the COVID-19 pandemic in spring 2020. However, a modified form of this prototype was tested in one of the groups. Figure 3 shows the steps in developing the rubric.

Data collection tools

In developing our assessment instrument, we made use of a number of qualitative data collection tools (Loiselle & Harvey, 2007). During rubric testing, the researcher, who was present in the classroom, recorded her observations in a journal. The teachers filled in self-evaluation forms on their use of the tool. These forms included questions on the various

Figure 3
Development-research stages

Phase 1 2018-2019 Fall 2019	Phase 2 Winter 2020	Phase 3 Spring 2020* 2020-2021
Design/implementation Prototype 1	**Design/implementation** Prototype 2	**Design/implementation** Prototype 3
Testing: LES 1 Marie class 1a Raphaëlle class 1a LES 2 Marie class 1a Raphaëlle class 1a LES 3 (fall 2019) Marie class 1b Raphaëlle class 1b Alice class 1	**Testing:** LES 1 Marie class 1b **Continued testing** Prototype 1 LES 1 Alice class 1 LES 2 Raphaëlle class 1b	**Testing:** Custom LES Alice class 2
Validation 1	**Validation 2**	**Validation 3** * Development interrupted due to COVID-19

sections of the learning and evaluating situations (LESs) and on the teachers' practices, for example, "I was able to assess the competency *Writes a variety of texts* with my students." They also took part in two or three semi-structured interviews that addressed their use of the rubric. These interviews included 13 questions, such as "Can you describe the professional acts you perform at the interpretation stage of the assessment process?" The interviews were systematically transcribed for analysis. At each stage of testing, we were able to record several pieces of information about the relevance of the rubric. It was on the basis of this information that specific changes were made to each validation phase.

Data analysis strategy

The data collected was analyzed over the course of the research, to ensure a complete understanding of the effectiveness of the prototypes used. An initial analysis using the metarubric determined the rubric's relevance. A score was assigned to each criterion in the rubric, providing an overall assessment of the prototype. We also conducted a thematic content analysis (Van der Maren, 2004) of the qualitative material from the semi-structured interviews and the self-evaluation forms, in order to identify the main themes that could guide us in making adjustments to the rubric. All of this was supported by observations recorded in the principal researcher's journal. These analyses were carried out using QDA Miner software (version 5.0.32) and were verified and accepted by the panel of experts involved in validating the rubric prototypes (Brantlinger et al., 2005). Thus, when the analysis identified an essential change to be made, the team discussed it beforehand. These discussions covered the adjustments needed to evolve the prototypes. Each development phase was rigorously examined and helped thoroughly document the development of our assessment instrument for French (language of Instruction) in middle school students with ASD.

Results

The results obtained are presented according to the development phases of the rubrics. For each phase, we present an analysis of the prototype, supported by examples and participants' perceptions. Although the process was carried out for the rubrics of the three competencies of the academic subject, for the purposes of this article, we present only excerpts for the competency *Reads a variety of texts*.

Phase I

The rubrics were initially developed to support specific learning activities for students with ASD. These activities were designed while incorporating major adjustments to the learning content. To this end, the research team had to redefine academic competencies in line with the QEP (MELS, 2006) and develop pedagogical activities including a teacher's guide and all the tools and materials needed for students to carry out the activities. The assessment instruments included rubrics for the different competencies targeted in French, as well as self- and peer-assessment activities.

The use of lower requirements, i.e., major content adjustments for competencies, made it possible to offer students expectations that were more within reach. For example, a set of competency elements for French (language of instruction) reduced the level of complication, while maintaining the level of complexity of the learning activities. This version, called "entry points," was used to design assessment instruments in line with the learning activities. The entry points were written upstream of all the teaching materials, in order to have a developmental continuum for academic competencies, with simpler levels of development enabling certain students to reach their zone of proximal development. The research team drew on the work of the Massachusetts Department of Elementary and Secondary Education (2017) to design an adapted continuum that would be consistent with the QEP (MELS, 2006). The learning tasks and assessment instrumentation were designed with these more accessible expectations in mind. Several versions of the rubrics were produced, validated and tested during the first design phase. The first versions used a vocabulary taken more from the entry points and the adapted continuum. To illustrate these initial attempts, we present an excerpt from the rubric for the competency Reads a variety of texts as an example, in Table 4.

The first versions of the rubrics were loaded with elements that weighed down the tool, including superfluous information. The columns "modified evaluation criterion" and "manifestation of criterion" and the headings "competency element," "entry points" and "access skills" proved unnecessary. To make the instrument more accessible, we consulted participants to better understand their students' needs and their level of competency development. Classroom observations recorded in the principal researcher's journal also made it possible to adjust the rubric and propose levels more representative of the realities experienced in the classroom. Table 5 shows an excerpt from the improved version of the rubric at the end of design/testing during phase 1.

This first prototype was tested by two teachers in spring 2019 and then by all three teachers in fall 2019. Raphaëlle and Alice readily recognized their students' levels in this rubric prototype, and found it easy to use. Indeed, the rubric helped draw an accurate picture of the students' competency levels, while still being in line with the QEP (MELS, 2006), and responded to the stated needs of the teachers:

> The rubric really gives you an overall picture, and I find it's a really nice wrap-up, and I don't have to go looking for it in my stuff [instructional tools and ministry documentation]. (Alice, interview 1)

> Right, the link with the QEP, which I don't really make, it does connect things, it streamlines the competency, it is precise, and it's very spelled out afterwards. At different levels, in terms of the link with French, a link is more clearly made with the QEP. (Raphaëlle, interview 1)

Marie had students with very heterogeneous academic profiles, some of whom did not have access to major accommodations, but rather to minor accommodations. In this case, the descriptors did not correspond to expectations for some students. The scale did not feature a continuum sufficiently broad to situate all the portraits of the students in this class, causing interpretation problems for the teacher. Moreover, she found the rubric difficult to use. It was not a tool Marie was used to using, and she struggled to understand its usefulness:

> When I read the rubrics, you know, I find the difference so slight between each [descriptor]. I have to go find the word that doesn't appear in the statement in order to be sure (Marie, interview 1).

In this passage, Marie identifies a major obstacle related to clarity of vocabulary, i.e., criterion 2 of the metarubric. Despite the close attention paid to words when writing descriptors, the complexity of the competency and the diversity of student profiles make the writing exercise complicated. It becomes unrealistic to achieve an unequivocal understanding of the instrument. However, by using simple vocabulary and discriminating words, one can aspire to craft a rubric that will be understandable to most recipients.

According to the criteria of the metarubric, this prototype seems relevant to two teachers, and "needs work" for the third. This explains why Raphaëlle and Alice continued to use this first prototype in winter 2020. For Marie, we agreed to add levels of performance to the rubric in order to take into account a wider variety of learners. We also reviewed the descriptors and tried to make them more distinctive. Table 6 sums up the changes that were made.

The accommodations put in place during the task were not addressed in this phase of development. Instead, the teachers focused on the competencies they had to teach and, above all, on taking ownership of the teaching materials. Only in phase 3 did we begin to reflect on the introduction of these accommodations.

Table 4

Excerpt from a rubric for the competency <u>Reads a variety of texts</u> (under development)

LES 1: Chocolate is so good! **Competency: Reads a variety of texts**							
Evaluation criterion	**Modified evaluation criterion**	**Competency element targeted (intermediate)**	**Manifestation of the criterion**	**Entry points**			**Access skills**
				A	**B**	**C**	**D**
Criterion 2 Sound interpretation*	The student, usually in their own words, formulates an interpretation from the texts and commits a few mistakes.	The student makes links between prior knowledge and new knowledge in relation to chocolate making.	Content	Student shares their new knowledge of chocolate making.	Student identifies new words in the text related to chocolate making.	Student identifies images that are new to them (that do not come from the text) and that represent chocolate making.	Student identifies an image related to chocolate making when presented with an image out of context.

*Translated freely from the French *interprétation fondée*, literally "founded interpretation."

Table 5

Excerpt from of prototype 1 for assessing the competency <u>Reads a variety of texts</u>

LES 3: Film critique **Competency: Reads a variety of texts**							
Evaluation criterion	**LES evaluation criterion**	**A** **Student…**	**B** **Student…**	**C** **Student…**	**D** **Student…**	**E** **Student…**	**F** **Student…**
Criterion 2 Sound interpretation	Quality of explanations Activity 5	gives clear, relevant and specific explanations regarding their appreciation of the film.	gives somewhat clear explanations.	gives a few explanations when asked by the teacher.	gives an overall comment when asked by the teacher.	points to the "like" group and gives an overall comment when asked by the teacher.	points to what they preferred (from among the elements they liked) by matching them to a positive motif (illustration).

Table 6
Results of analysis pertaining to protype 1 testing

Participants	Overall appraisal	Aspects to work on
Raphaëlle and Alice	B – Relevant, but the accommodations provided were not addressed.	C7 – Document adjustments made to assessment.
Marie	D- Redo.	C2 – Clarify vocabulary C3 – Limited choice of criteria prevents a vision of the entire competency. C4 – Descriptors do not allow an evaluation of a student with minor accommodations C5 – Insufficient no. of levels to be representative of all students. C7 – Document adjustments made to the assessment.

Phase 2

The second prototype presents a rubric to which levels of performance have been added to cover the diversity of learner development, using descriptors that are more in line with QEP expectations. However, as the learning tasks were developed for students with cognitive challenges, it was difficult to write these new descriptors, given that the expected productions remained simple. For this reason, we tried to design a rubric that would be closer to the instrumentation used in regular classrooms, which represented an additional challenge. Table 7 shows an excerpt from this rubric.

The intent of this second prototype was to include all student portraits in a group within a situation where some were entitled to major accommodations and others were not. We quickly realized that we could not have middle school-level expectations with simple tasks. As we tried to make the expectations more complex, descriptors were moving further and further away from the learning students were working on in the classroom. What is more, adding levels of performance made the rubric more cumbersome, which hampered analysis of competency development. Table 8 shows the elements to be worked on in this second prototype.

Table 7

Excerpt from Prototype 2 for assessing the competency <u>Reads a variety of texts</u>

SAÉ 1 : Chocolat que c'est bon ! **Compétence : Lire des textes variés**									
Evaluation criterion	**LES evaluation criterion**	**3** **The student…**	**2** **The student…**	**1** **The student…**	**0** **The student…**	**00** **The student…**	**A** **The student…**	**B** **The student…**	**C** **The student…**
Criterion 2 Sound interpretation	Quality of links made	makes many judicious links between key elements on chocolate making.	makes relevant links between several key elements of chocolate making and what the student learned.	makes links between general elements pertaining to chocolate making and what the student learned.	makes simple links between key elements of chocolate making and what the student learned	makes generally correct links between key elements of chocolate making.	makes links between prior and new knowledge of chocolate making.	makes links with the teacher's help between prior and new knowledge about chocolate making.	shares their new knowledge of chocolate making.

Table 8
Results of analysis pertaining to experimentation with Prototype 2

Participant	Overall appraisal	Aspects to improve
Marie	C – needs work	Cl – The rubric is overloaded, including too many levels. C4 – The descriptors poorly illustrate what is worked on in class. C6 – The descriptors are appropriate to the task but the task is not appropriate for all students. C7 – Accommodations are included in the descriptors.

In this rubric, as in Prototype 1, we included the help provided by the teacher in the descriptors. For example, level of performance B contains: "makes connections, with the teacher's help, between prior and new knowledge [...]". We discussed this at length with the participants, as they noted that some students could demonstrate a lower level of competence, but independently (e.g., the student makes links between prior and new knowledge), while others demonstrated higher levels of competence, but with help (e.g., student makes relevant links between several key elements). Our descriptors were too hermetic, confusing the teachers' judgment, especially when it came to scoring.

> Yes, it's going to be modified because I'm still here. You know, that's clear for me. If I'm here, if I remove questions. You know, I sort of know the limits, but to put it between 60 and 100, because that's our report cards, where do I put it? (Marie, interview 1)

At the end of this second validation phase, we noted that including a large number of descriptors in a single rubric was inefficient. When the level of expectation varies greatly from one student to the next, it becomes essential to present different tasks, according to the students' needs. This also calls for different instruments to take this into account. In the context of our project, the proposed LESs included major accommodations, particularly in terms of content and complexity. For example, it was difficult to observe a Secondary 1 [first year of middle school] level for a competency. In addition, the fact that the teacher's help was included in the descriptors created a barrier to using the instrument.

Phase 3

We designed our third prototype in the spring of 2020, during COVID-19 pandemic lockdowns. The team took the time to analyze the various data collected since the beginning of the experiment and came to the following conclusions: It was preferable to return to a smaller number of levels of performance, specifically written for students with modified expectations, and to distinguish between the student's level of autonomy and their level of mastery of the content. Indeed, separating the assistance provided from the academic content made it easier to document the type of accommodation made. Were the accommodations production-related (form, content, level of complexity, etc.) or support-related (technological assistance, teacher assistance, etc.)? Table 9 shows an excerpt from the third prototype.

This prototype could not be tested in the same contexts as the previous prototypes. Of the three participants, only Alice was in service during the 2020-2021 year, and her classroom group required only minor accommodations, unlike the students initially concerned by our material. In view of our observations the previous year, we advised the teacher not to use the material developed by the SRSE. That said, with Alice's participation, we took the time to co-construct a new rubric based on her needs and in alignment with the principles learned during the experiment. The first rubric was developed to assess the competency to write texts at the middle school level, including all the evaluation criteria. We wrote task-specific descriptors and added a space for the teacher to indicate any minor adjustments applied during the activity. Table 10 shows an excerpt from this rubric, which the teacher found helpful.

> Including the various adjustments [accommodations] or specific comments on the actual help or support given to the student during the task enables a better shared understanding of the student's level of autonomy in relation to a school task at their grade level. (Alice, reflection sheet, custom LES)

The third phase of rubric validation was interrupted by the COVID-19 pandemic. We were unable to consolidate testing of this prototype with the SRSE instructional materials, but the information obtained from the third teacher was encouraging.

Table 9

Excerpt from Prototype 3 for assessing the competency <u>Reads a variety of texts</u>

LES: Chocolate is so good! **Competency: Reads a variety of texts** **Level of support: No support (NS)**		**Occasional support (OS)**		**High support (HS)**			
Evaluation criterion	**LES evaluation criterion**	**A**	**B**	**C**	**D**	**E**	**F**
Criterion 2 Sound interpretation	Quality of links made	Student makes links between prior and new knowledge of chocolate making.	Student discusses their new knowledge of chocolate making and makes links with their knowledge.	Student discusses their new knowledge of chocolate making, but some are off topic.	Student identifies new words in the text on chocolate making.	Student points to images that are new to them (that are not from the text) and that represent chocolate making.	Student points to an image pertaining to chocolate making when presented with an image that is out of context.
		NS OS HS	NS OS HS	NS OS HS	NS OS HS	NS OS HS	NS OS HS

Table 10

Excerpt from Prototype 3 adapted to Alice's situation to assess the competency <u>Reads a variety of texts</u>

		A	**B**	**C**	**D**	**E**
		The student...	**The student...**	**The student...**	**The student...**	**The student...**
Adaptation to writing situation	Respect du sujet	describes several aspects of a character including physical characteristics and personality using well-developed ideas.	describes physical and psychological characteristics of a character using well-developed ideas.	summarily describes physical and psychological characteristics of a character; ideas are vague.	briefly describes physical and psychological characteristics, gives no specifics.	describes disjointed aspects or restates ideas without reformulating them in their own words.
	Minor adjustments made					

The following discussion outlines the successes and areas for improvement identified during our development process. In addition, we will offer an overall analysis of the participants' use of the rubrics, and guidelines for creating rubrics for special education students.

Discussion

The aim of our study was to develop a specific rubric for French (language of instruction) competencies (reading, writing, oral communication) for middle school students with ASD.

The initial findings concern the heterogeneity of learner profiles, despite the common diagnosis of ASD (Poirier et al., 2017). Within the same group of students, the profiles and needs vary, which supports the need for a criterion-based interpretation of the evolution of academic competencies for these students, i.e., comparing productions to criteria and not one production to another (Simon & Forgette-Giroux, 2001). However, creating a specific assessment instrument remains an arduous process. An important limitation is the flexibility of the instrument, as identified by Goodrich (1996). Contrary to what this author points out, too much heterogeneity among learners becomes a major obstacle to designing a descriptive rubric. To pinpoint the difficulties encountered, the systematic use of the metarubric proved helpful. This instrument enabled us to quickly identify and synthesize the elements to be worked on when evaluating the relevance of our prototypes.

First of all, we reworked the relevance and number of levels of performance, as well as the suitability of content in the first two prototypes. As Leroux and Mastracci (2015) point out, having too many levels of performance hinders an effective assessment of the level of competency. It is preferable to maintain a limited number of levels of performance and to adjust the tasks required of students so that they can demonstrate their level of competence. To achieve this, it is necessary to establish differentiation mechanisms that respect students' pace of development, by making major adjustments during learning and assessment.

Next, the clarity of vocabulary was an element that needed to be worked on. Indeed, Marie clearly mentioned that the vocabulary used in the rubrics could be difficult to understand: it was therefore essential to choose a simple vocabulary that could be understood by the majority of

recipients (e.g., parents and students themselves). On this subject, Wiertz et al. (2020) point out that a joint construction of the tool enables this understanding, which we also observed during the development of the last rubric. Indeed, since Alice had participated in this development, it was much easier for her to use and thoroughly understand the descriptors.

An important element of our rubric is the distinct identification of the accommodations put in place. As Leroux and Mastracci (2015) observed, it is preferable not to include accommodations in the descriptors of a rubric. However, leaving some free space for more personalized writing of accommodations would also be a favourable element to add. Alice seems to have appreciated this approach. For others, it would be a place to leave comments.

At the end of this analysis, we identified three key principles to consider when creating rubrics for subject-specific competencies in students with ASD:

- Maintain a reasonable number of descriptors for a given task. To this end, the creation of subgroups within the same class would make it possible to assess the competency levels of students along a similar continuum;
- Use simple, comprehensible vocabulary. To achieve this, co-construction with teachers seems to be helpful;
- Distinguish between adjustments and descriptions of observable events. In a context of major accommodations, the descriptors are already proof of the proposed adjustments, but it is also interesting to document the level of support provided to students. In the case of minor accommodations it is a good idea to leave a space reserved for identifying adjustments or for comments.

Conclusion

While our study presents an innovative look at rubrics, certain limitations should be noted. First, the interruption in the development process prevented us from carrying out a final comprehensive analysis of the tool. In addition, the main tests were carried out during the same school year and with a small sample. It would therefore be interesting to take the principles identified in our study and experiment with them in a new context that includes students with ASD, by testing over a fairly long period, with

external support provided by a specialist or by an educational advisor. It would also have been appropriate to include teachers in the development of the rubric from the outset. Considering how successful this was with Alice at the end of the project, it could have been beneficial for our entire approach.

The rubric is an interesting instrument for judging the level of competency of students with ASD. However, on several occasions, a participant raised the following point:

> You know, if you gave me two weeks, based on my eight students, I'd draw up fantastic portraits for you. I'd use the learning progression to situate them and identify what they've achieved. I'd do just that and I'd have zero trouble telling if they were in Secondary 1 or Modified Secondary 1. But I don't have that kind of time (Marie, interview 1).

This teacher indicated that she wanted to make an even more descriptive interpretation than that allowed by the rubric, i.e., a narrative assessment (Margrain, 2010). The other participants also raised this point, saying they were more comfortable writing comments to communicate their students' level of competence. As a result, a tool such as a portfolio could enhance the effectiveness of the rubric (Derycke, 2000). Teaching to students with ASD is a context for which assessment practices are still underdeveloped (Lamarche & Durand, 2021), and this is also true for all types of students across the special education sector (Branciard et al., 2016). In a context where the Quebec school system values the inclusion of students in regular classes, it is essential to deepen our knowledge of assessment practices for students requiring minor or major accommodations, in order to ensure fair, equal and equitable assessment for all students.

Submitted: December 20, 2022
Final version: July 19, 2023
Accepted: August 26, 2023

REFERENCES

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]. Cognitio. 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]. Cognitio. 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 1re 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/feduc.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

Calhoon, 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 5e 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 (2e 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. Raîche, 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), 130144. https://doi.org/10.1016/j.edurev.2007.05.002

Jung, L. A., & Guskey, T. R. (2011). *Grading Exceptional and Struggling Learners*. Corwin.

Lamarche, 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

Louis, R. (2004). L'évaluation des apprentissages, théorie et pratique (2e 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., & Arsenault, 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 (16e 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/prfrmsec1ercyclev2.pdf

Ministère de l'Éducation, du Loisir et du Sport (MELS). (2011). *Cadre d'évaluation des apprentissages : Français, langue d'enseignement, enseignement secondaire 1er et 2e cycle*. Gouvernement du Québec. http://www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/formation_jeunes/francais-langue-d-enseignement-sec.pdf b

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*, 129144. 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 6e 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 (2e ed.). Stylus Publishing, LLC.

Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2006). *Classroom assessment for student learning. Doing it right - using it well* (2e 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* (2e 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