

Validation of the French version of the classroom assessment scoring system infant and toddler in Quebec

Lise Lemay¹, Paul Hayotte², Julie Lemire¹, Caroline Bouchard³, Lorie-Marlène Brault-Foisy¹,
Nathalie Bigras¹, Gilles Cantin¹

¹Department of Didactics, Faculty of Education, Université du Québec à Montréal, Montréal, Canada

²Department of Psychology, Faculty of Social Sciences, Université du Québec à Montréal, Montréal, Canada

³Department of Teaching and Learning Studies, Faculty of Education, Université Laval, Québec, Canada

Article Info

Article history:

Received Aug 26, 2022

Revised Nov 20, 2023

Accepted Nov 29, 2023

Keywords:

Classroom assessment scoring
Early childhood education
Measurement tools
Psychometric properties
Quality of interactions

ABSTRACT

The objective of this study is to validate the French version of the Classroom Assessment Scoring System (CLASS) infant and toddler, as employed to assess the quality of interactions in groups of children under 3 years old, in childcare centers in Quebec where French is the official language. Indeed, when using a different language version of a standard-based tool outside its original context, an important step is to verify that it remains reliable and valid for measuring the research construct. This validation study was conducted in Montreal area (Quebec, Canada). The subjects were 154 classrooms (46 infant, 108 toddler) located within a representative sample of 68 childcare centers. Live classroom observations were conducted in the fall 2018 with the CLASS and other measures of process quality. Results replicate the factor structures of the original versions of the CLASS tool and provide evidence for the good reliability (inter-rater reliability, internal consistency) and validity (criterion and construct) of the French versions. The discussion highlights cross-cultural differences in the classrooms, childcare centers, and regulations that could explain some differences obtained in this research and, therefore, needs to be considered when using the CLASS in French to have a reliable and valid tool to measure the quality of interactions.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Lise Lemay

Department of Didactics, Faculty of Education, Université du Québec à Montréal

C.P. 8888, Succursale Centre-Ville, Montréal (Québec), Canada

Email: lemay.lise@uqam.ca

1. INTRODUCTION

In early childhood education, the quality of educator-child interactions can be considered a determining factor of educational quality for children's development and learning [1]–[3]. However, research shows that children aged 0 to 3 are being exposed to interactions whose quality can be considered low to moderate [4], [5]. Based on these findings, it seems vital to take steps to improve the quality of interactions between educators and children under the age of 3 in childcare centers.

To that end, an increasing number of governments and countries are implementing quality assessment and improvement systems [6]. In Quebec, the Ministère de la Famille (Ministry of the Family) is hoping to use the assessment and improvement of educational quality of educational childcare service facilities (EAQS) measure to ensure monitoring of the quality provided to children, especially those under 3 years of age. To do so, interaction quality will be measured using the Classroom Assessment Scoring System (CLASS), as in other places around the world.

The CLASS Pre-K (Précolaire in French) [7], designed for observation in groups of children aged 3 to 5, has been widely adopted for assessing interaction quality in recent years, both in the United States and worldwide [8]. It has been rapidly adopted in practice, research and monitoring largely due to the demonstrated quality of its psychometric properties [9], [10]. More recently, the tool was adapted for observing groups of younger children, including an infant tool (6 weeks to 18 months) [11] and a toddler tool (15 to 18 months) [12]. Although warm, sensitive, and stimulating interactions are considered crucial in all versions of the CLASS tool, their complexity grows in proportion to the level of development of the children. As shown in Figure 1, the greater number of dimensions and domains shows the various manifestations of this according to the child's age and the version of the CLASS [13]. The CLASS infant and toddler are used increasingly widely to assess the quality of interactions in groups of children under 3 years of age [14], including in systems focused on quality assessment and improvement across the globe and particularly in the United States (US) [15].

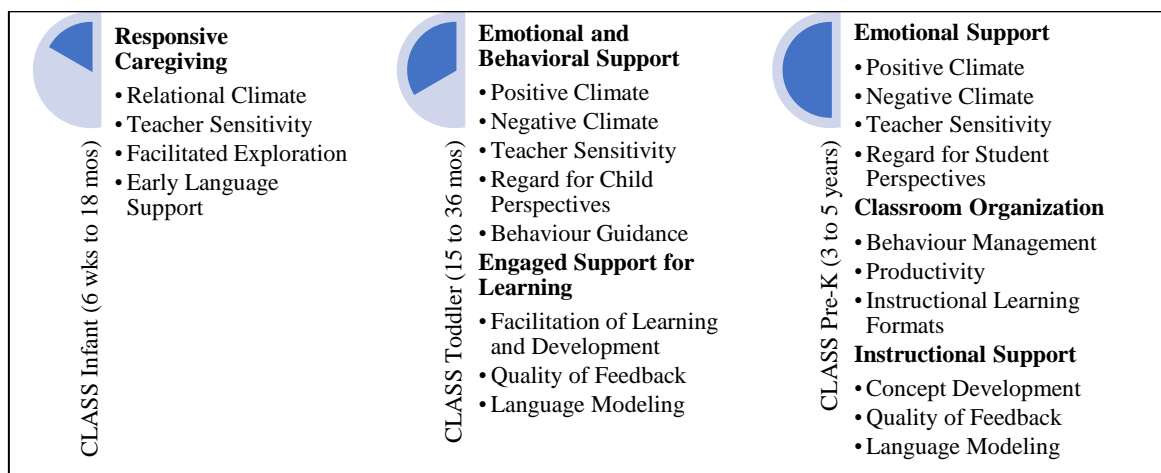


Figure 1. Dimensions and domains of the CLASS, infant, toddler, and Pre-K

To use the CLASS infant (“Poupon” in French) and CLASS toddler (“Trottineur” in French) in Quebec for the Ministry’s EAQS measure, we must first validate the tool when used in another socio-educational context as well as in another language such as French, the official language of Quebec. Certain issues need to be addressed using an observational tool outside the context in which it was developed [8]. For example, certain studies conducted in Australia [16], Quebec [17], France [18], Chile [19], and Finland [20] were unable to replicate the factor structure of the original version of the CLASS Pre-K. In fact, researchers from Finland [20], Portugal [21], and Sweden [22] found that certain dimensions of the CLASS did not seem to apply very well to their socio-educational context. Such findings underscore the importance of validating whether the French version of the CLASS preserves the psychometric properties that the original version is famed for, to make sure that we are using a rigorous tool to assess the quality of interactions to which children aged 0 to 3 are exposed in childcare centers. In other words, are the French versions of the CLASS infant and toddler demonstrate good reliability and validity when used to measure quality of interactions in Quebec, a new linguistic and socio-educational context where they need to be implemented? This study answers that question.

In Quebec (Canada), educational childcare services (ECS) have seen dramatic growth following changes to the province’s family policy in 1997. With these changes, the government created a network of regulated childcare services, some of which were subsidized and accessible with a reduced family contribution (\$8.70 per day in 2022), while others were non-subsidized and available at the full rate (around \$40 per day). According to the latest data, in a census of 521,952 Quebec children under 6 years of age [23], 285,407 spots were available in childcare services. The majority of those spots were in early childhood center facilities (CPEs), which are not for profit subsidized facilities (CPEs=100,664 spots), in subsidized private daycares (50,103 spots) and non-subsidized private daycares (68,302 spots) and in subsidized home childcare services (66,338 spots) [24]. All of these providers have a legal obligation to offer an educational program that is consistent with basic principles of partnership with parents, active and play-based learning, the uniqueness of every child and the child’s holistic development [25].

Since 2017, all providers are also required to participate in a process to assess and improve educational quality in order to ensure that they are providing educational quality that supports children's educational success and holistic development [26]. That year, the Quebec Ministry of the Family revised the law to add that requirement regarding the EAQS measure. The measure was implemented in stages beginning in 2019, with the objective of assessing educational quality in groups of children aged 3 to 5 in CPEs and daycares. Work in groups of children aged 3 years and under, the subject of this article, was undertaken in 2017–2018. The EAQS measure identifies the qualities of educator-child interactions, assessed using CLASS, as the main measurement of educational quality.

The CLASS infant and toddler are more recent and less documented than the CLASS Pre-K. The studies published to date, many of them from the United States, indicate that the psychometric properties of CLASS infant and toddler seem to be satisfactory when used in English [4], [11], [27]–[33]. More specifically, what has been documented to date are the factor structure, reliability (inter rater, internal consistency) and validity (criterion and construct) of each original version of the tool.

For the CLASS infant [11], the US data show a factor structure in a single domain, responsive caregiving, which is used to assess interactions that demonstrate the adult's response to children's emotional and cognitive needs. However, the model's adjustment values were significantly improved by allowing correlation between the dimensions labelled as emotional (relational climate and teacher sensitivity) and those labelled as educational (facilitated exploration and early language support) [11]. Additionally, one pilot study conducted in Belgium suggested that a two-factor structure could potentially be a reasonable adjustment, with one domain concerning emotional support and the other educational support [34].

For CLASS toddler, previous study [12] have suggested a two domain factor structure to address the emotional and behavioral support the adult provides and the engaged support for learning their interactions give to children. Several other studies have agreed about this structure [12], [27], [29], [34]–[36]. One study conducted in Portugal and Finland likewise supported this two domain structure, after excluding the dimension negative climate [37]. Dutch researchers, however, have suggested a three-factor structure for the CLASS toddler, separating emotional support from behavioral support to make two domains, with engaged support for learning as the third [32]. Lastly, a US study pointed to the need to examine the factor structure of the CLASS toddler when used in other linguistic and cultural contexts [35].

Two types of reliability are generally reported for the CLASS: inter rater reliability and internal consistency. For the CLASS infant, a US study found an absolute agreement rate of 88.00% [30] between observers, and a Portuguese study found one of 99.00% [38], both suggesting strong inter-rater reliability. According to the instrument's authors, the inter rater reliability of the CLASS toddler can also be judged to be "good" with a mean agreement rate of 83.00% [12]. European studies showed an absolute observer agreement rate of 82.14% in the Netherlands and 92.75% in Poland [39], as well as 98.00% in Finland and 96.00% in Portugal [37]. In terms of internal consistency, often measured with Cronbach's alpha (α), data collected with the CLASS infant had internal consistency coefficients from "acceptable" to "excellent" in the domain of Responsive Caregiving ($\alpha=0.73$ – 0.97) [11]. The US authors likewise confirmed the "high" internal coherence of the two domains of CLASS toddler [27].

More specifically, for the emotional and behavioral support domain, internal consistency coefficients were $\alpha=0.77$ in the Netherlands [39], $\alpha=0.86$ in Finland, $\alpha=0.88$ in Poland [39], as well as $\alpha=0.94$ [36], and $\alpha=0.95$ [37] in Portugal. For the Engaged Support for Learning domain, the authors reported coefficients of $\alpha=0.87$ in Finland [37] and $\alpha=0.92$ in Portugal, Poland and the Netherlands [36], [37], [39]. A Swiss study, however, reported lower internal consistency coefficients, with $\alpha=0.69$ for the emotional and behavioral support domain (rising to $\alpha=0.79$ after exclusion of the negative climate dimension) and $\alpha=0.80$ for engaged support for learning [40].

For the criterion validity, the CLASS scores are associated with contextual variables. CLASS scores are correlated with ratio and teacher's experience in the United States [30], as well as teacher training, group size and ECS location in Portugal [28] for the infant version. On the toddler version used in American and Dutch studies, CLASS scores are related with teacher variables, such as years of experience, specialized training, depressive symptoms, or beliefs [27], [29], [32], group variables such as size and ratio [32], [33], and ECS variables such as type of setting [32].

The construct validity of the CLASS infant and toddler has been established by comparison to another process quality observational tool, the Infant/Toddler Environment Rating Scale, Revised (ITERS-R) [41]. Previous studies showed that the scores obtained in the CLASS correlate to the process quality in ECSs for the subscales of "Interaction" and "Listening and Talking," which are also constructs measured by the CLASS (convergent validity). However, it correlates less with ECSs structure, including the subscales "Space and Furnishings" and "Program Structure," which are not measured by the CLASS (divergent validity) [11].

Overall, although US, Portuguese, Dutch and Belgian studies have documented the psychometric properties of the original versions of the CLASS infant and toddler, more research (particularly research

conducted in other linguistic and socio-educational contexts) is needed for several reasons. Firstly, because there does not seem to be a consensus on certain properties, such as the factor structure. Secondly, because validation data come primarily from US studies or studies that use the instrument in its original English language versions. This study, therefore, aims to contribute to the existing understanding of the factor structure, reliability, and validity of the CLASS infant and toddler, and to do so by using French versions of the tool in Quebec. The main objective of this study is to document the psychometric properties of the French versions of the CLASS infant and toddler (that is, CLASS Poupon and Trotteineur), as employed to assess the quality of interactions in groups of children under 3 years of age, in ECS facilities in Quebec. Within that, it aims to achieve the following specific objectives for each French version of the CLASS: i) document the factor structure; ii) evaluate the reliability in terms of inter rater reliability and internal consistency; iii) test the criterion and construct validity.

2. RESEARCH METHOD

2.1. Recruitment and sample

Data for this study was collected in the greater Montréal area (Quebec, Canada), where most of the ECS facilities are located, throughout fall 2018. A total of 408 ECS facilities were selected randomly from a list provided by Quebec's Ministry of the Family and contacted about participation. To be eligible for the study, an ECS facility had to care for children under the age of 3, include at least two groups of children to be observed and two teachers who agreed to participate in the project, have been operating for four years or more and use French to communicate.

Selection criteria were set to achieve the representativeness of the sample. First, the study was looking to have proportional representation of the three types of ECS facilities in Quebec (31% CPEs, 22% subsidized daycares and 32% non-subsidized daycares). We also hoped to include ECSs showing a variety of socioeconomic circumstances (30% facilities located in disadvantaged areas), which were identified using Quebec's Material and Social Deprivation Index [42]. Another recruitment goal was for 10% of ECS facilities to be ones using a specific educational approach other than Quebec's *Accueillir la petite enfance* program (e.g., Montessori, High/Scope).

A total of 68 ECS facilities met the selection criteria and agreed, voluntarily and without compensation, to participate in the study. These facilities provided a sample size of 154 groups, of which 46 were observed with the CLASS Poupon (infant) and 108 with the CLASS Trotteineur (toddler). This includes a representative sampling of the different types of ECSs, of facilities located in disadvantaged neighborhoods and of specific educational approaches.

2.2. Procedures and instruments

The research team contacted the administrator of each of the 68 ECS facilities to plan an observation day. On that day, a team of observers went to each facility and spent the same five hours collecting data in participating groups using observation instruments or interviews. They first assessed the quality of interactions using the CLASS Poupon or the CLASS Trotteineur. Next, to confirm construct validity for the purposes of this validation study, they assessed other aspects of educational quality: the quality of teacher-child relationships; the quality of the physical environment; and the quality of child observation and planning practices. A second observer was present in the group for 17% of observations and interviews to calculate inter rater reliability.

2.2.1. Quality of educator-child interactions

As previously stated, the quality of interactions between educators and children was assessed using the CLASS Poupon (groups of children aged 6 weeks to 15 months) and Trotteineur (groups of children aged 15 to 36 months). The French language versions of the CLASS tool, CLASS Poupon and Trotteineur, are translations of the original English language instruments. The translations were done with authorization from Teachstone, the company responsible for the CLASS tool in the United States. Translating the written material and producing the video material into French adapted to the Quebec ECSs context was a multi-step process: i) initial translation of the material from English to French; ii) revision by a bilingual review committee and cross-cultural adaptation; iii) validation of the translation by a group of experts. All of these steps were necessary to ensure the quality of the written and video material in French [43], and are laid out in more detail to provide an overview of the research team's process.

a. Translation, revision, and validation of written material

The written components of the CLASS—PowerPoint presentation, training participant guide, manual, score sheets—were translated from English to French by a bilingual professional translator. The

CLASS Trottineur was translated first, and the CLASS Poupon second. Particular attention was given to translating concepts and to a transcultural adaptation consistent with the concepts used in early education in Quebec, rather than a strictly literal translation. For each French version of the CLASS, a group of three bilingual experts in early childhood education reviewed and validated the tool's initial translation. Where the experts were not in consensus about some elements, additional measures were taken. Certain translated words and troublesome formulations were researched in the Quebec literature on early childhood education to confirm that the translation was using the appropriate terminology for a diverse range of concepts.

Next, research professionals in early childhood education gave the validated initial versions a final review to confirm the quality of the content, accuracy of the language and consistency of terminology used between the CLASS Poupon and Trottineur. Finally, a bilingual collaborator from Teachstone read over the CLASS Poupon and Trottineur and provided commentary to ensure that the translations respected the intentions of the original instrument. A few minor edits were made in response to this external review to finalize the French versions of the instrument.

b. Production, revision, and validation of video material

To be consistent with Teachstone's original training material, the research team also needed to produce the video material required for training activities and certification of French-speaking observers. Short videos of one to two minutes apiece, exemplar videos, were produced, with eight videos in total for CLASS Poupon and 14 for CLASS Trottineur. Additionally, for each tool, 10 longer videos of 15 to 20 minutes apiece were filmed with the appropriate age group: five videos for training exercises in scoring and five videos for the CLASS certification.

Next, a committee of certified observers, experienced with educational observation and the CLASS Poupon and Trottineur, assigned scores to each of these videos (long videos). At the end of this work, they were able to set expert judges' scores (master codes) to be used in determining the reliability of observers in scoring exercises and certification. The research team also created written content to accompany the videos (descriptions of the exemplar videos, master code justifications).

Throughout the process, the production of all material was checked, commented on, and approved by Teachstone representatives. This means that Teachstone staff approved the content of the training and certification videos, the master code scores given by the expert judges in Quebec, the texts accompanying the exemplar videos and the master code justifications. This work was carried out from October 2017 to July 2018 and, once it was complete, Teachstone authorized the research team to provide training sessions on the French language versions of the CLASS tool (Poupon and Trottineur versions) and to conduct a study documenting the psychometric properties of the tools used in Quebec.

The next phase was to conduct a study of the psychometric properties of the French versions of the CLASS tools [43]. During the data collection, the CLASS Poupon and Trottineur follow roughly similar procedures for use. The Poupon assessments were done in 25-minute cycles, and the Trottineur assessments in 30-minute cycles. At the end of each cycle, the observer rated each of the dimensions by giving scores on a 7-point scale (1 or 2=low; 3, 4 or 5=moderate; 6 or 7=high). For both tools, a minimum of four observation cycles (a total of one hour and 40 minutes for the CLASS Poupon and two hours for the CLASS Trottineur) are required for the data to be considered valid by the instrument authors. At the end of the assessment, the scores of the four cycles are averaged, and the scores for the instrument's dimensions and domains are the resulting means.

In this study, the observers (n=20) were trained and certified on the CLASS Poupon and Trottineur over four days of training and two additional days of certification. Following Teachstone's procedure, to be certified to use the tool, the observers had to obtain an 80% agreement rate with the master codes when scoring the five long videos. A difference of one point between the observers' scores and the master codes was tolerated when determining agreement. Observers were also required to demonstrate their ability to reliably observe each dimension of the tool, with no more than three instances of disagreement for a given dimension. This procedure corresponds exactly with the one developed by Teachstone for the original tool.

All CLASS training participants received their certification, most of them on the first try. Although the threshold required for passing was 80%, the mean agreement rate was 90.94% for the CLASS Poupon and 89.55% for the CLASS Trottineur. These results, which reflect an agreement rate of roughly 9 scores out of 10, are an indication of good inter-rater reliability. It was also agreed that measures should be put in place to document the reliability of all training and certification material to ensure that no bias was introduced during translation. After participating in the French language training for the CLASS Poupon and Trottineur, therefore, five bilingual observers (25% of the observer team) who had not been familiar with the CLASS before their training were invited to complete the online certification offered in English by Teachstone. They scored five different videos in English. All five observers successfully obtained certification in English, with a mean agreement rate of 86.25% for CLASS infant and 85.63% for CLASS toddler.

These data have two implications. On the one hand, they allowed us to validate the material translated and developed by the research team. On the other hand, they suggested that the written and video material in French reliably prepared the observers to use the CLASS for assessments in Quebec, since the observers had no difficulty in obtaining CLASS certification in both the French versions and the original English versions.

2.2.2. Caregiver-child relationships

The research used the Caregiver Interaction Scale (CIS) [44], which rates the quality of the relationship between children and a caregiver, to assess the construct validity (convergent and divergent) with the CLASS tool. The CIS is composed of 23 items distributed among three subscales: sensitivity, harshness, and detachment. Each item is given a score between 1 and 4, with a higher number meaning a greater frequency of a given behavior except in the case of those items with inverted phrasing.

2.2.3. Quality of the physical environment

The quality of the physical environment was observed in under 30 minutes using 11 items to assess whether ECS rooms were welcoming, flexible, allowed a diverse range of activities and groupings, were tailored to the needs of children and educators, encouraged independence [45], [46]. Based on the number of elements checked according to the scoring guide, each item was given a score on a scale from 1 (minimum) to 4 (very good). The scores were then reported on a scale of 1 to 7 to be consistent with other instrument use. The mean of the item scores provided the total score for quality of room arrangements.

2.2.4. Quality of child observation and planning practices

The quality of child observation and planning practices was documented using a 30-minute semi-structured interview with each educator of an observed group [47]. This interview included a requirement that the observer review the documents the educator stated they used to plan and to observe children, to validate their responses [48], [49]. Based on the interview responses, the observer used the scoring guide to rate the eight sections of the interview as low, moderate, or high quality. Then, the observer used the ratings for the first four sections to determine a score for the quality of child observation practices on a scale of 1 (low quality, all sections rated as low quality) to 7 (high quality, all sections rated as high quality). They did the same for the last four sections of the interview to determine a planning quality score. Finally, the quality scores for child observation and planning practices were averaged to obtain an overall quality score.

3. RESULTS AND DISCUSSION

3.1. Results

Table 1 (CLASS Poupon) and Table 2 (CLASS Trotteur) give the means and standard deviations found in the study, presented in parallel with the data from the validation study of the original version of the tool. Descriptive statistics show that the scores obtained using the French versions of the CLASS are similar, although generally higher, than the scores obtained using the original English versions of the tool. These scores range from moderate to high quality for the relational climate and teacher sensitivity dimensions in CLASS Poupon and for the dimensions in the emotional and behavioral support domain in CLASS Trotteur, and from low to moderate for the facilitated exploration and early language support dimensions in CLASS Poupon and the dimensions in the engaged support for learning domain in CLASS Trotteur.

Table 1. Descriptive statistics of the CLASS Poupon's domain and dimensions for the pilot project and in Teachstone's reference data

	Pilot research n=46			Teachstone (CDR)* n=56			Teachstone (UNC)** n=97		
	M	SD	Range	M	SD	Range	M	SD	Range
Relational climate	5.65	0.89	2.25-7.00	4.11	1.23	1.00-6.00	5.04	1.02	2.00-6.67
Teacher sensitivity	5.52	1.02	2.00-7.00	4.03	1.28	1.00-6.00	4.75	1.16	1.50-7.00
Facilitated exploration	3.92	1.36	1.00-7.00	3.38	1.18	1.00-6.00	3.68	1.07	1.00-6.50
Early language support	4.11	1.38	1.00-7.00	3.23	1.08	1.00-6.00	3.31	1.19	1.00-6.25
Responsive caregiving	4.80	1.05	1.56-6.75	4.02	n/a	n/a	4.20	n/a	n/a

*Child Development Resources (CDR) is a research project conducted in Virginia between June 2012 and March 2013 in 56 infant classrooms within childcare centers and home-based childcares.

**University of North Carolina (UNC) conducted a research project between 2012 and 2013 in 97 infant classrooms within childcare centers.

Table 2. Descriptive statistics of the CLASS Trotteiner's domains and dimensions for the pilot project and in Teachstone's reference data

	Pilot research n=108			Teachstone (NCR-LAP)* n=93		
	M	SD	Range	M	SD	Range
Positive climate	5.48	1.04	2.25-7.00	5.03	1.22	1.75-7.00
Negative climate	1.36	0.55	1.00-5.00	2.70	0.90	1.00-6.25
Teacher sensitivity	5.13	1.10	1.50-7.00	4.33	1.16	2.00-7.00
Regard for child perspective	4.26	1.27	1.50-7.00	4.36	1.05	1.50-6.75
Behavior guidance	4.92	0.96	2.50-6.75	4.07	1.29	1.75-6.50
Facilitation of learning and development	3.44	1.01	1.50-6.75	3.43	1.20	1.00-6.50
Quality of feedback	2.94	1.04	1.00-6.00	n/a	n/a	n/a
Language modeling	3.28	1.07	1.25-6.50	2.22	1.07	1.00-5.00
Emotional and behavioral support	5.29	0.84	2.60-6.95	4.62	n/a	n/a
Engage support for learning	3.22	0.97	1.25-6.42	2.83	n/a	n/a

*The North Carolina Rated License Assessment Program (NCR-LAP) is a study conducted between 2009 and 2010 at University of North Carolina at Greensboro

3.1.1. Factorial structure

To verify that the data collected was a good match for the factorial structure of the original version of the CLASS, we used Mplus Version 8 [50] to conduct confirmatory factor analysis (CFA) on each French versions of CLASS. To assess the two models' fit for the data, we used the following fit statistics: Chi-squared (χ^2), comparative fit index (CFI), root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR). In addition to the goodness of fit test of differences between the models (χ^2), which is sensitive to sample size, we were able to use several other statistics reported in the technical appendices of the CLASS manual to test the model fit. The CFI compares the proposed model to another model, with the assumption that there is no relationship between the variables. A CFI>0.90 is considered a good fit [51], [52]. The RMSEA estimates the extent to which the model is an acceptable fit for the data; for a good fit, the value should be under 0.06 [53], [54]. The SRMR measures the discrepancy between the model and the data; the lower the SRMR, the better the fit, and an SRMR<.08 is considered "acceptable" [52]. Changes in the CFI, RMSEA and SRMR were reviewed to compare the models, but they were not compared against standards.

a. CLASS Poupon

Table 3 presents the CFA results for the CLASS Poupon. These results show that a single factor solution does not have "good" fit indexes based on the criteria (RMSEA (90% CI)=0.572 (0.416–0.745); $\chi^2(2)$ =35.41, p <0.001; CFI=0.830; SRMR=0.091). The model's goodness of fit improves significantly, however, when the model allows for correlation between the dimensions Relational Climate and Teacher Sensitivity, and between the dimensions facilitated exploration and early language support (RMSEA (90% CI)=0.160 (0.000–0.443); $\chi^2(1)$ =2.30, p <0.129; CFI=0.993; SRMR=0.011). Although the fit indexes were not at the level of a "good" fit, the results for the criteria assessed suggest that the data collected do indeed replicate the structure in the single domain of the CLASS infant.

Table 3. Confirmatory factor analysis for the CLASS Poupon's domain for the pilot project and in Teachstone's reference data

	Pilot research		Teachstone	
	CFA without cross-loadings	CFA with cross-loadings	CFA without cross-loadings	CFA with cross-loadings
Relational climate	0.98	0.79	0.90	0.69
Teacher sensitivity	0.92	0.69	0.95	0.75
Facilitated exploration	0.78	0.86	0.75	0.92
Early language support	0.74	0.89	0.73	0.88

Since the authors of the CLASS Infant conducted analyses in which correlation between the dimensions Relational Climate and Teacher Sensitivity and between the dimensions Facilitated Exploration and Early Language Support was permitted, a second CFA was carried out for exploratory purposes to confirm the hypothesis of a structure in two domains. Table 4 shows the results. When the residual variance of the Relational Climate dimension is set to zero, the fit indexes support a two domain solution (RMSEA (90% CI)=0.072 (0.000–0.297); $\chi^2(2)$ =2.53, p <0.283; CFI=0.997; SRMR=0.010), with one domain containing the dimensions relational climate and teacher sensitivity and the other containing the dimensions facilitated exploration and early language support.

Table 4. Confirmatory factor analysis in two domains for the CLASS Poupon

Dimensions	Standardized coefficients	
Relational climate	1.00	
Teacher sensitivity	0.90	
Facilitated exploration		0.96
Early language support		0.90

b. CLASS Trotlineur

Table 5 gives the CFA results for the CLASS Trotlineur. Although the fit indexes for the two domain structure of CLASS Trotlineur are not optimal based on the criteria (RMSEA (90% CI)=0.080 (0.043–0.115); $\chi^2(19)=36.20$, $p<0.010$; CFI=0.974; SRMR=0.056), they are nonetheless considered to be “acceptable.” This suggests that the data replicate the two domain structure of the original CLASS Toddler.

Table 5. Confirmatory factor analysis for the CLASS Trotlineur’s domains for the pilot project and in Teachstone’s reference data

	Pilot research		Teachstone	
	Standardized coefficients		Standardized coefficients	
Positive climate	0.88		0.89	
Negative climate	0.56		0.58	
Teacher sensitivity	0.88		0.95	
Regard for child perspective	0.76		0.86	
Behavior guidance	0.88		0.88	
Facilitation of learning and development		0.88		0.95
Quality of feedback		0.92		n/a
Language modeling		0.88		0.80

3.1.2. Reliability

a. Inter rater reliability

Inter rater reliability measures verify that observers attribute similar dimensional scores when observing the same situation. Out of 26 observations made to judge inter rater reliability, the average absolute agreement rate among observers during data collection was 86% for both the CLASS Poupon and the CLASS Trotlineur. These values have been deemed to be entirely “satisfactory” [55].

b. Internal consistency

Results of our analysis of the internal consistency of the CLASS tools yielded Cronbach’s alpha coefficient values of $\alpha=0.91$ for the Responsive Caregiving domain of the CLASS Poupon. For the CLASS Trotlineur, the Cronbach’s alpha coefficient values were $\alpha=0.89$ for the emotional and behavioral support domain and $\alpha=0.92$ for the engaged support for learning domain. These values are considered entirely “satisfactory”.

3.1.3. Validity

a. Criterion validity

Univariate analysis of variance (ANOVAs) was done to verify whether the scores in the CLASS Poupon and Trotlineur were associated with certain contextual variables: i) Type of educational childcare service (ECS); ii) Material and Social Deprivation Index (MSDI); and ii) Use of a specific educational approach. An initial analysis confirmed that the data satisfied the premises of ANOVAs. Table 6 shows the results.

Table 6. Univariate ANOVA for the domains of the CLASS Poupon and Trotlineur

	Poupon		Trotlineur			
	Responsive caregiving		Emotional and behavioral support		Engaged support for learning	
	F	Effect size	F	Effect size	F	Effect size
Type of ECS	13.97***	0.39	19.95***	0.28	15.78***	0.23
Deprivation index	0.32	0.04	1.39	0.06	1.78	0.08
Use of a specific educational approach	15.18***	0.26	9.38**	0.08	21.66***	0.17

Note: * $p<0.05$; ** $p<0.01$; *** $p<0.001$

The ANOVAs showed several significant differences depending on the type of ECS. For the CLASS Poupon, scores in the responsive caregiving domain were significantly higher in CPEs than in subsidized daycares ($p=0.25$; 95% CI=0.087, 1.669) and non-subsidized daycares ($p=0.00$; 95% CI=0.778, 2.227). For the CLASS Trottineur, the scores were likewise significantly higher in CPEs than in daycares, for both the domains: emotional and behavioral support (subsidized daycares: $p=0.00$; 95% CI=0.473, 1.339; non-subsidized daycares: $p=0.00$; 95% CI=0.448, 1.257) and engaged support for learning (subsidized daycares: $p=0.00$; 95% CI=0.439, 1.468; non-subsidized daycares: $p=0.00$; 95% CI=0.423, 1.384). According to the values [56], these differences are “large” ones, with variance explained of, respectively, 39.39%, 27.54% and 23.11%. There was no significant difference between subsidized and non-subsidized daycares.

We did not find any significant differences in the domain scores for the CLASS Poupon and Trottineur based on an ECS’s Material and Social Deprivation Index (MSDI). When an ECS uses a specific educational approach, however, the results do show a significant difference in CLASS scores. Compared to educators working in ECSs that stated that they implement the *Accueillir la petite enfance* educational program [25], those in ECSs that practice another educational approach received higher scores in the CLASS Poupon’s Responsive Caregiving domain ($F 1,45=15.18$, $p=.000$) and in the CLASS Trottineur’s emotional and behavioral support domain ($F 1,107=9.38$, $p=.003$) and engaged support for learning domain ($F 1,107=21.66$, $p=.000$). These significant differences are considered to be “moderate” (emotional and behavioral support) or “large” in size (responsive caregiving, engaged support for learning).

b. Construct validity

To verify the construct validity, correlation analysis was conducted between the measures of interaction quality, caregiver-child relationship (CIS), quality of the physical environment and quality of child observation and planning practices. Table 7 shows the results of this correlation analysis. According to Cohen’s values [56], moderate to strong correlations were noted among all the instruments used to measure educational quality. Strong positive correlations were found between the CIS Sensitivity scale and the responsive caregiving domain of the CLASS Poupon ($r=0.65$) and the emotional and behavioral support ($r=0.60$) and engaged support for learning ($r=0.52$) domains of the CLASS Trottineur. Moderate correlations were found between the domains of the CLASS Poupon and Trottineur and the measure of quality of the physical environment and the CLASS domains ($r=0.41-0.46$), and the measure of quality of child observation and planning practices ($r=0.48-0.67$). Lastly, negative correlations were found between the domains of the CLASS Poupon and Trottineur and the CIS Harshness ($r=-0.36- -0.51$) and Detachment ($r=-0.42- -0.63$) scales.

Table 7. Correlations between the domains of the CLASS Poupon and Trottineur with other dimensions of process quality

	CLASS Poupon		CLASS Trottineur
	Responsive caregiving	Emotional and behavioral support	Engage support for learning
Physical environment	0.42**	0.41***	0.46**
Observation and planning practices	0.67***	0.48***	0.42***
Caregiver-child relationship sensitivity	0.65***	0.60***	0.52***
Caregiver-child relationship harshness	-0.36*	-0.51***	-0.44***
Caregiver-child relationship detachment	-0.63***	-0.46***	-0.42***

Note. * $p<.05$; ** $p<.01$; *** $p<.001$

3.2. Discussion

The research attests to the importance of educator-child interaction quality in ECSs for infant and toddler development [2]. At the same time, modest quality levels are reported internationally for children under 3 years of age [4], [5], [28], [40]. These findings highlight the importance of monitoring and improving the quality of interactions with children aged 0 to 3, as the Quebec Ministry of the Family hopes to do. The overall objective of this study, then, is to contribute to making available a rigorous measurement tool for assessing the quality of educator-child interactions in ECSs by documenting the psychometric properties of the CLASS Poupon and the CLASS Trottineur as used in French in groups of children under 3 years of age in Quebec ECS facilities.

The results of this study show that practically all the scores given with the CLASS Poupon and Trottineur are higher than the validation scores in the original versions of the tool, except for the CLASS Trottineur dimensions Regard for Child Perspectives and Facilitation of Learning and Development, where scores were lower. Researchers in Belgium and the Netherlands likewise reported higher scores in ECSs [32], [34]. Similar differences were previously reported with the CLASS Pre-K, with higher scores given in Quebec than in France [57] or in the United States with CLASS Pre-K, including for emotional support [58].

Without access to the reported validation data, however, we cannot conclusively determine whether the score differences between countries are significant. Although the score values for the CLASS Poupon and Trottineur were generally slightly higher, the data follow similar trends to the US reference data collected with the original versions. The difference between the higher quality interaction levels found for emotional and behavioral support interventions and the lower ones for engaged support for learning correspond to patterns generally observed in research worldwide [30], [31].

Although the score data collected for the CLASS Poupon and Trottineur show certain similarities to the original versions' score data, then, they are also distinct in certain ways, which may be due to the study's cultural contexts. For example, in their validation study in Finland, Pakarinen *et al.* [20] reported higher CLASS Pre-K scores than those given in the United States. They discuss these differences in terms of process quality, pointing out that the educational practices of teachers in Finnish kindergartens are characterized by the absence of a negative climate, the use of child initiated activities and small group activities consistent with developmentally appropriate practices, and in terms of structural quality, such as teachers who have received specialized training in early childhood education, lower ratios and smaller group sizes and the age of the children in the groups (older than in US kindergarten). These are interaction quality and structural quality characteristics that can also describe the socio-educational context of Quebec ECSs [57].

In a similar vein, in their report on the educational quality of ECSs in Portugal, England, the Netherlands, and Germany, Slot *et al.* [59] showed that the relationships between structural quality and process quality variables were not the same across the countries they studied. The authors concluded by pointing to the complex interrelations between multiple structural quality variables of the early childhood educational system of each country to explain the corresponding process quality they observed. Furthermore, the OECD [60] stated that 40% of the countries and territories surveyed in its report did not have shared pedagogical frameworks for care settings for children aged 5 and under, which likely influenced the quality of interactions observed. Although sociocultural differences may have been present, it seems very relevant to follow up on the data collected with the CLASS tools, both English or French versions, in order to better understand these results and detect any potential persisting anomalies in the scores for various CLASS dimensions [43].

For the purposes in assessing the psychometric properties of the CLASS Poupon and Trottineur in their French versions, three areas of this study are fruitful to discuss. The research will address the factor structure of the scales; their reliability; and their validity. The findings in each of these areas will be discussed in turn, including the socio-educational differences that emerge.

3.2.1. Factor structure

Our findings for the CLASS Poupon allowed us to replicate the one domain structure developed for the original instrument [11], but the model fit indexes were not rated as "good" until residuals had been correlated. On this basis, after exploring this modification and the two domain structure discussed by previous researchers [34], our factor analysis showed that a two domain structure for the CLASS Poupon would allow for better model fit. As in the same authors' study in Belgium, one domain would include dimensions relational climate and teacher sensitivity (privileging caregiving) and the other domain would include dimensions facilitated exploration and early language support (privileging education).

These findings may attest to cultural differences between the United States and other countries where the CLASS is used, as mentioned in earlier report [59]. This could include parental leave, which lasts for 4 to 20 months in Belgium and 18 months in Quebec, meaning that children begin attending ECSs later than in the United States, where the average length of parental leave is 3 months. More recent data indicate that 50% of Quebec children start to regularly attend an ECS at the age of 12 months [61]. For more precise figures, in March 2017, regulated ECSs received 2.3% of children under 6 months of age, 6.5% of children aged 6 to 11 months, 8.9% aged 12 to 17 months, 11.1% aged 18 to 23 months and 21.7% aged 24 to 36 months [62]. Therefore, children's daily routine in the United States and Quebec seems to differ, based on the fact that the younger children are, the more time seems to be given on a daily basis to routines and transitions [59]. Some researchers have also suggested the routines are largely underutilized as learning opportunities. Since children are on average older in the groups observed in Quebec, they are closer to the age of toddlers (which begins at 15 months for CLASS Trottineur). From this perspective, in groups of infants in Quebec, the fact that the children are older could mean that less time is spent on daily care routines and that educator infant interactions could be more complex, which could be an influence leading the factor structure of CLASS Poupon to be more similar to that of the CLASS Trottineur. Further research will be needed to confirm the hypothesized two domain structure for the CLASS Poupon.

Our findings for the CLASS Trottineur allowed us to replicate the two domain structure developed for the original instrument. Furthermore, the coefficients we found were relatively like those for the original English version [12], [31], [32], [35]. However, the model fit indexes were not ideal, so we had three options:

i) reject the model; ii) choose the best fit among those suggested by the literature; or iii) modify them based on fit measures and diagnostic statistics. Of these options, it seems to us that a three-domain structure would be fruitful to explore with a larger sample in a future validation phase [32].

3.2.2. Reliability

One significant challenge that is inherent to observing interaction quality is training observers to use the tool's framework. Generally speaking, the reliability of the CLASS Poupon and Trottineur is satisfactory. The inter rater reliability demonstrates that the observers gave similar scores for the same situation, and even did so more consistently than the data of other studies [12], [30]. Some authors have pointed to the role of training and calibration procedures for inter rater reliability with the CLASS [63]. The training procedures used and the research project's guidance after every instance of inter rater agreement seem to have had a positive effect on the observers' ability to reliably observe and a limiting effect on the potential for individual deviation post certification.

As regards the second type of reliability assessed in this study, the internal consistency of the domains of the CLASS Poupon and the CLASS Trottineur was very good or excellent. The internal coherence of the CLASS Poupon's domain was very similar to that observed in other studies conducted with the instrument [30]. The internal coherence of CLASS Trottineur's domains is similar to the values reported in studies in Finland [37] and Poland [39]. The CLASS Poupon and Trottineur, in French, seem equipped to allow accurate (homogenous) and consistent scoring.

3.2.3. Validity

The data suggests that this tool does successfully measure interaction quality. Our findings also allow us to flesh out the criterion validity of the CLASS Poupon and Trottineur, beyond the group variables (e.g., ratio, size) and educator variables (e.g., training, depressive symptoms) explored in prior research [14], [27]–[30], [33]. Only Barros *et al.* [28] looked at ECS variables. Our findings show that ECS variables such as facility type and use of a specific educational approach are connected to all domains of interaction quality.

The quality of interactions measured with the CLASS Poupon and Trottineur varies by type of ECS observed. This variation by types of ECS is also found for measures of global educational quality [64]–[67] of quality of interactions measured with CLASS Pre-K [68] and of quality of interactions measured with CLASS Toddler [32], [59]. These results are consistent with the aforementioned studies with regard to the distribution of quality levels among various types of ECS, with higher levels of quality in not-for-profit childcare services (CPEs) compared to commercial childcare services (daycares).

The quality of interactions also varies depending on the use of specific educational approaches, a finding that was not significant [32]. Edwards [69] found that an educational approach includes five major elements supporting quality: interactions, content, routines, activities and resources. It is possible that, in settings that use a specific educational approach, collective reflection within the facility and associated guidance may increase pedagogic intentionality for the whole team, including with regard to the five elements that support quality [69]. It is also possible that such a process may be particularly beneficial in groups of infants and toddlers for whom, as we have seen, routines and care activities take up more time in the daily schedule even though those contexts are underutilized for interactions that support learning. There is as yet little data available on the topic of use of specific educational approaches in ECSs [70], so future research would help us understand these findings more fully.

The study also found that socioeconomic characteristics were not connected to the CLASS domains. One potential explanation could be the type of measurement used for these characteristics, which were determined using the postal code of the facility where the ECS was located. This measure may be more distal than a measure of the people (adults and children) in the setting. Such a proximal measure would allow for more in-depth assessment of this issue, since socioeconomic conditions were associated with educational quality in previous studies where it was measured based on whether or not the ECS had received a grant for disadvantaged areas [67], [71].

Lastly, significant correlations (moderate to strong) between the quality of interaction measures with the CLASS and other dimensions of process quality suggest that the CLASS Poupon and Trottineur have good convergent and divergent validity. Groups with higher educator child interaction quality had higher scores for quality of the physical environment as measured by observation scale, for quality of child observation and planning practices as assessed via interviews, and for sensitivity of caregiver child relationships, as well as lower scores for the Harshness and Detachment scales of CIS. It is possible, therefore, to state that there are strong positive correlations with the CIS Sensitivity scale, the closest interaction constructs to that measured by the CLASS, and weaker correlations with the quality of the physical environment. This finding is consistent with those reported with the ITERS-R [11], [12]. The findings also agree with those of Jamison *et al.* [30], who stated that the CLASS is an observational measurement of interactions rather than physical environment.

3.2.4. Study limitations

One limitation is that, when interpreting our findings, it is important to note that this study had a small sample of 154 groups in ECS that volunteered to participate in a preliminary study phase. An upcoming phase would be to carry out a broader study with a sample that is more representative of the population. Also, the data collected with the CLASS Poupon and Trotteineur were not putted in relation with the children's development to assess validity. Although such relationships have been reported in previous studies [12], [27], [30], it would nonetheless be fruitful to document the predictive validity of the French versions of the CLASS, including infant and toddler development measures. The findings of this study must be interpreted with nuance, particularly if they are used to inform policies or practices for a wider population.

4. CONCLUSION

The importance of high quality early childhood education, in combination with research pointing to the modest quality levels of interactions with infants and toddlers, attests to the relevance of monitoring, maintaining and improving the quality of interactions with children under 3 years of age in ECSs. This study has found that the psychometric properties of two versions of the CLASS tool for that purpose, the CLASS Poupon and the CLASS Trotteineur, are satisfactory. These rigorous instruments, in their French versions, provide information about the quality levels of interactions between educators and children aged 0 to 3 years in Quebec ECSs. However, to achieve similar good psychometric properties, this paper highlights ECS variables that need to be considered when implementing the CLASS in French and in any other socio-educational context. In addition, more studies are needed to deepen cross-cultural differences in the classrooms, childcare centers, and regulations associated with the scores on the CLASS Poupon and Trotteineur in larger samples. Still, data collected with these French versions of CLASS could inform public policies for assessing and improving educational quality to ensure that the youngest members of our society can develop to their fullest potential to give equal opportunities to all.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the Québec Ministry of the Family and the Université du Québec à Montréal for contributing financially to this study. Additionally, our warmest gratitude goes to the childcare educators and managers who invested their valuable time by participating in the study, as well as to the research professionals and assistants who gathered the data. We would like to thank our valuable colleagues, members of the Quality of early childhood education research team, who provided helpful feedback on a previous version of this manuscript.

REFERENCES




- [1] M. C. Araujo, M. Dormal, and N. Schady, "Childcare quality and child development," *Journal of Human Resources*, vol. 54, no. 3, pp. 656–682, 2019.
- [2] J. Cadima, G. Nata, S. Barros, V. Coelho, and C. Barata, "Literature review on early childhood education and care for children under the age of 3," *OECD Education Working Papers*, No. 243, OECD Publishing, Paris, 2020, doi: 10.1787/a9cef727-en.
- [3] T. J. Sabol, S. Soliday Hong, R. C. Pianta, and M. R. Burchinal, "Can rating pre-K programs predict children's learning," *Science*, vol. 341, no. 6148, pp. 845–846, 2013.
- [4] S. Barros and C. Aguiar, "Assessing the quality of Portuguese child care programs for toddlers," *Early Childhood Research Quarterly*, vol. 25, no. 4, pp. 527–535, 2010.
- [5] M. Pessanha *et al.*, "Stability and change in teacher-infant interaction quality over time," *Early Childhood Research Quarterly*, vol. 40, pp. 87–97, 2017.
- [6] Organisation for Economic Co-operation and Development (OECD), *Starting strong IV: Monitoring quality in early childhood education and care*. OECD, 2015.
- [7] R. C. Pianta, K. M. La Paro, and B. K. Hamre, *Classroom Assessment Scoring System® (CLASS®) Pre-K*. Baltimore: Brookes, 2008.
- [8] L. Lemay, J. Lehrer, and M. Naud, "Le CLASS pour mesurer la qualité des interactions en contextes culturels variés," (in French), *Les dossiers des sciences de l'éducation*, no. 37, pp. 15–34, 2017.
- [9] N. Bigras, J. Lemire, and J. Eryasa, "Comparaison d'instruments d'observation de la qualité de services éducatifs de la petite enfance," (in French), *Les dossiers des sciences de l'éducation*, no. 37, pp. 35–51, 2017.
- [10] H. Li, J. Liu, and C. V. Hunter, "A Meta-Analysis of the Factor Structure of the Classroom Assessment Scoring System (CLASS)," *The Journal of Experimental Education*, vol. 88, no. 2, pp. 265–287, 2020, doi: 10.1080/00220973.2018.1551184.
- [11] B. K. Hamre, K. M. La Paro, R. C. Pianta, and J. LoCasale-Crouch, *Classroom Assessment Scoring System® (CLASS™) manual, Infant*. Baltimore: Paul H. Brookes Publishing Co., 2014.
- [12] K. M. La Paro, B. K. Hamre, and R. Pianta, *Classroom Assessment Scoring System® (CLASS®) Toddler*. Baltimore: Brookes, 2012.
- [13] G. Joseph, E. Feldman, C. Brennan, R. Naslund, J. Phillips, and A. Petras, "Seeds to Success field test: Year two-final technical report," University of Washington: Center for Research and Training, Childcare Quality and Early Learning, vol. 11, 2011.

- [14] P. L. Slot, "Structural characteristics and process quality in early childhood education and care: A literature review," *OECD Education Working Papers*, No. 176, OECD Publishing, Paris, 2018, doi: 10.1787/edaf3793-en.
- [15] National Center on Early Childhood Quality Assurance, "QRIS resource guide," 2017. [Online]. Available: <https://ecquality.acf.hhs.gov/resource-guide>.
- [16] D. Cloney, C. Nguyen, R. J. Adams, C. Tayler, G. Cleveland, and K. Thorpe, "Psychometric Properties of the Classroom Assessment Scoring System (Pre-K): Implications for Measuring Interaction Quality in Diverse Early Childhood Settings," *Journal of Applied Measurement*, vol. 18, no. 3, pp. 299–318, 2017.
- [17] C. Bouchard, G. Cantin, A. Charron, H. Crépeau, and J. Lemire, "Validation du Classroom Assessment Scoring System (CLASS) en maternelle 4 ans à mi-temps au Québec," (in French), 82e congrès de l'ACFAS, 2014.
- [18] P. Dessus, O. Cosnefroy, and G. Joët, "Evaluer la qualité des interactions enseignant-élèves en début de scolarité: Qualités psychométriques," (in French), 82e congrès de l'ACFAS, 2014.
- [19] E. Treviño *et al.*, "Teacher practices and learning improvement in Chilean preschool classrooms," in *The annual meeting of the 56th Annual Conference of the Comparative and International Education Society*, San Juan, Puerto Rico, 2012.
- [20] E. Pakarinen *et al.*, "A validation of the Classroom Assessment Scoring System in Finnish kindergartens," *Early Education and Development*, vol. 21, no. 1, pp. 95–124, 2010.
- [21] J. Cadima, C. Peixoto, and T. Leal, "Observed classroom quality in first grade: associations with teacher, classroom, and school characteristics," *European Journal of Psychology of Education*, vol. 29, no. 1, pp. 139–158, 2014, doi: 10.1007/s10212-013-0191-4.
- [22] M. Norling, A. Sandberg, and L. Almqvist, "Engagement and emergent literacy practices in Swedish preschools," *European Early Childhood Education Research Journal*, vol. 23, no. 5, pp. 619–634, 2015, doi: 10.1080/1350293X.2014.996423.
- [23] Observatoire des tout-petits, *Comment se portent les tout-petits au Québec? Portrait 2021*. Montréal, Québec: Fondation Lucie et André Chagnon (in French), 2021.
- [24] Ministry of Family, "Number of childcare centres and licensed spaces at March 31, 2022." [Online]. Available: http://www.mfa.gouv.qc.ca/fr/publication/Documents/places_0.pdf (accessed: Jun. 6 2022).
- [25] Ministry of Family, *Meeting Early Childhood Needs. Educational program. For educational childcare services*. Québec: Les publications du Québec (in French), 2019.
- [26] Ministry of Employment and Social Solidarity. *Educational Childcare Act*. [Online]. Available: <https://www.legisquebec.gouv.qc.ca/en/document/cs/S-4.1.1>
- [27] E. Bandel, N. Aikens, C. A. Vogel, K. Boller, and L. Murphy, "Observed quality and psychometric properties of the CLASS-T in the Early Head Start Family and Child Experiences Survey," Mathematica Policy Research, 2014.
- [28] S. Barros *et al.*, "Infant child care quality in Portugal: Associations with structural characteristics," *Early Childhood Research Quarterly*, vol. 37, pp. 118–130, 2016.
- [29] S. Castle, A. C. Williamson, E. Young, J. Stubblefield, D. Laurin, and N. Pearce, "Teacher–child interactions in early head start classrooms: Associations with teacher characteristics," *Early Education and Development*, vol. 27, no. 2, pp. 259–274, 2016.
- [30] K. R. Jamison, S. Q. Cabell, J. LoCasale-Crouch, B. K. Hamre, and R. C. Pianta, "CLASS–Infant: An Observational Measure for Assessing Teacher–Infant Interactions in Center-Based Child Care," *Early Education and Development*, vol. 25, no. 4, pp. 553–572, 2014, doi: 10.1080/10409289.2013.822239.
- [31] K. M. La Paro, A. C. Williamson, and B. Hatfield, "Assessing quality in toddler classrooms using the CLASS-Toddler and the ITERS-R," *Early Education and Development*, vol. 25, no. 6, pp. 875–893, 2014.
- [32] P. L. Slot, J. Boom, J. Verhagen, and P. P. Leseman, "Measurement properties of the CLASS Toddler in ECEC in The Netherlands," *Journal of Applied Developmental Psychology*, vol. 48, pp. 79–91, 2017.
- [33] A. C. Thomason and K. M. La Paro, "Measuring the quality of teacher–child interactions in toddler child care," *Early Education and Development*, vol. 20, no. 2, pp. 285–304, 2009.
- [34] M. Vandenbroeck, P. L. Slot, and H. Hulpia, "Quality in home-based childcare providers: variations in process quality," *European Early Childhood Education Research Journal*, vol. 29, no. 2, pp. 261–277, 2021.
- [35] K. Bichay-Awadalla and R. J. Bulotsky-Shearer, "Examining the Factor Structure of the Classroom Assessment Scoring System Toddler (CLASS-T) in Early Head Start and Subsidized Child Care Classrooms," *Early Education and Development*, vol. 33, no. 2, pp. 309–325, 2021.
- [36] C. Guedes, J. Cadima, T. Aguiar, C. Aguiar, and C. Barata, "Activity settings in toddler classrooms and quality of group and individual interactions," *Journal of Applied Developmental Psychology*, vol. 67, p. 101100, 2020, doi: <https://doi.org/10.1016/j.appdev.2019.101100>.
- [37] J. Salminen, C. Guedes, M.-K. Lerkkanen, E. Pakarinen, and J. Cadima, "Teacher–child interaction quality and children's self-regulation in toddler classrooms in Finland and Portugal," *Infant and Child Development*, vol. 30, no. 3, p. e2222, 2021, doi: <https://doi.org/10.1002/icd.2222>.
- [38] A. I. Pinto *et al.*, "Quality of infant child care and early infant development in Portuguese childcare centers," *Early Childhood Research Quarterly*, vol. 48, pp. 246–255, 2019.
- [39] O. Wysłowska and P. L. Slot, "Structural and process quality in early childhood education and care provisions in Poland and the Netherlands: A cross-national study using cluster analysis," *Early Education and Development*, vol. 31, no. 4, pp. 524–540, 2020.
- [40] T. Diebold and S. Perren, "The impact of childcare-group situational age composition on caregiver–child interactions," *European Journal of Developmental Psychology*, vol. 17, no. 4, pp. 598–615, 2020.
- [41] T. Harms, D. Cryer, and R. M. Clifford *Infant/Toddler Environment Rating Scale—Revised edition*. Teachers College Press, 2003.
- [42] P. Gamache, D. Hamel, and R. Pampalon, "L'indice de défavorisation matérielle et sociale: en bref," Institut national de santé publique du Québec, Montréal (in French), 2015.
- [43] B. A. Bracken and A. Barona, "State of the art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment," *School Psychology International*, vol. 12, no. 1-2, pp. 119–132, 1991.
- [44] J. Arnett, "Caregivers in day-care centers: Does training matter?" *Journal of Applied Developmental Psychology*, vol. 10, no. 4, pp. 541–552, 1989, doi: [http://dx.doi.org/10.1016/0193-3973\(89\)90026-9](http://dx.doi.org/10.1016/0193-3973(89)90026-9).
- [45] G. Cantin, L. Lemay, and J. Lemire, *Outil d'évaluation. Aménagement des lieux. Version pouponnière*. Montréal: Université du Québec à Montréal (in French), 2017.
- [46] G. Cantin and J. Lemire, *Outil d'évaluation. Aménagement des lieux. Version 18 mois ou plus*. Montréal: Université du Québec à Montréal (in French), 2017.
- [47] L. Lemay, G. Cantin, J. Lemire, and C. Bouchard, "Conception and validation of the quality of educators' observation and planning practices scale (QEOPPS)," *Early Years*, vol. 41, no. 2-3, pp. 144–160, 2021.





- [48] G. Cantin, L. Lemay, J. Lemire, and P. G. Hayotte, *Entrevue destinée à l'éducatrice. Qualité des pratiques d'observation des enfants et de planification expériences vécues par les enfants (qui sous-tendent les expériences vécues par les enfants - QPOP). Version pouponnière*. Montréal: Université du Québec à Montréal (in French), 2017.
- [49] G. Cantin and J. Lemire, *Entrevue destinée à l'éducatrice. Qualité des pratiques d'observation des enfants et de planification expériences vécues par les enfants (qui sous-tendent les expériences vécues par les enfants - QPOP). Version 18 mois et plus*. Montréal: Université du Québec à Montréal (in French), 2017.
- [50] L. K. Muthén and B. O. Muthén, *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén, 1998.
- [51] P. M. Bentler, "Comparative fit indexes in structural models," *Psychological Bulletin*, vol. 107, no. 2, p. 238, 1990.
- [52] L. t. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 6, no. 1, pp. 1–55, 1999.
- [53] M. W. Browne and R. Cudeck, "Alternative ways of assessing model fit," *Sociological Methods & Research*, vol. 21, no. 2, pp. 230–258, 1992.
- [54] J. H. Steiger, "Structural model evaluation and modification: An interval estimation approach," *Multivariate Behavioral Research*, vol. 25, no. 2, pp. 173–180, 1990.
- [55] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.
- [56] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale: Lawrence Erlbaum, 1988.
- [57] N. Bigras, P. Dessus, L. Lemay, C. Bouchard, and C. Lequette, "Qualité de l'accueil d'enfants de 3 ans en centre de la petite enfance au Québec et en maternelle en France," (in French), *Enfances Familles Générations. Revue interdisciplinaire sur la famille contemporaine*, no. 35, 2020.
- [58] M.-C. Bouchard, "Relation entre la qualité des interactions éducatrice-enfants dans le groupe en centre de la petite enfance et la qualité des interactions de l'enfant de quatre ans avec ses pairs," *Maîtrise*, Université Laval, Québec (in French), 2015.
- [59] P. L. Slot, P. P. Leseman, J. Verhagen, and H. Mulder, "Associations between structural quality aspects and process quality in Dutch early childhood education and care settings," *Early Childhood Research Quarterly*, vol. 33, pp. 64–76, 2015.
- [60] Organisation for Economic Co-operation and Development (OECD), *Starting Strong VI: Supporting Meaningful Interactions in Early Childhood Education and Care*. OECD, 2021.
- [61] A. Lavoie, L. Gingras, N. Audet, F. Lapointe, and B. Perron, *Enquête Québécoise sur le Parcours préscolaire des Enfants de Maternelle 2017*. Institut de la statistique du Québec (in French), 2019.
- [62] Ministry of Family. Situation of early childhood centers, daycares and family daycares in Quebec in 2017. Analysis of the 2016-2017 activity reports submitted by the divisions of educational childcare service companies. (in French). [Online] Available: https://www.mfa.gouv.qc.ca/fr/publication/Documents/Situation_des_CPE_et_des_garderies-2015.pdf
- [63] E. Westergård, S. K. Ertesvåg, and F. Rafaelsen, "A preliminary validity of the classroom assessment scoring system in Norwegian lower-secondary schools," *Scandinavian Journal of Educational Research*, vol. 63, no. 4, pp. 566–584, 2019.
- [64] C. Drouin, N. Bigras, C. Fournier, H. Desrosiers, and S. Bernard, *Grandir en qualité 2003. Enquête québécoise sur la qualité des services de garde éducatifs*. Montréal: Institut de la statistique du Québec (in French), 2004.
- [65] A. S. Fuligni, C. Howes, Y. Huang, S. S. Hong, and S. Lara-Cinisomo, "Activity settings and daily routines in preschool classrooms: Diverse experiences in early learning settings for low-income children," *Early Childhood Research Quarterly*, vol. 27, no. 2, pp. 198–209, 2012, doi: <http://dx.doi.org/10.1016/j.ecresq.2011.10.001>.
- [66] L. Gingras, A. Lavoie, and N. Audet, "Quebec' national study on child care quality - Growing up in quality 2014. Quality of not-for-profit child care centers," (in French), Institut de la statistique du Québec, Québec, 2015, vol. 2.
- [67] L. Gingras, A. Lavoie, and N. Audet, "Quebec' national study on child care quality - Growing up in quality 2014. Quality of for-profit unsubsidies child care centers," (in French), Institut de la statistique du Québec, Québec, 2015, vol. 3.
- [68] G. Cantin, J. Lemire, L. Lemay, P. G. Hayotte, and C. Bouchard, "Bilan du projet pilote "Évaluer et améliorer la qualité éducative dans les services de garde en installation". Livrable n°6. Rapport produit pour le ministère de la Famille," Université du Québec à Montréal, Montréal (in French), 2017.
- [69] S. Edwards, "Process quality, curriculum and pedagogy in early childhood education and care," OECD Education Working Paper No. 47, 2021.
- [70] Y. Anders, *Literature review on pedagogy*. Paris: Organisation for Economic Co-operation and Development, 2015.
- [71] C. Japel, R. E. Tremblay, and S. Côté, "Quality Counts! Assessing the Quality of Daycare Services Based on the Quebec Longitudinal Study of Child Development," (in French), *Choix*, vol. 11, no. 4, 2005.

BIOGRAPHIES OF AUTHORS







Lise Lemay    is a full professor in early childhood education didactics at Université du Québec à Montréal (Québec, Canada). She is a member of the "Qualité des contextes éducatifs de la petite enfance" research team. Her interests are early childhood education quality, home-based childcare, play-based approach, make-believe play and children' development. She can be contacted at email: lemay.lise@uqam.ca.







Paul Hayotte     is a graduate student associated with the “Qualité des contextes éducatifs de la petite enfance” research team. His research focuses on the well-being of grandparents in Quebec, and the contribution of new technologies in maintaining intergenerational relationships. He can be contacted at email: hayotte.paul_georges@courrier.uqam.ca.







Julie Lemire     is a research professional in early childhood education at UQAM’s Department of Didactics since 2007. She holds a bachelor’s degree in child psychology, a master’s degree in neuropsychology, and a DESS in higher education pedagogy. She is part of the research team that contributed to the adaptation of the CLASS tools into French for Quebec and to the development of the Ministry of the Family “Assessment and improvement of educational quality of educational childcare service facilities” measure. She can be contacted at email: Lemire.julie@uqam.ca.







Caroline Bouchard     is a full professor at Laval University (Quebec, Canada). Her work focuses on the quality of interactions and support for adults who work with young children. She is a regular researcher in the “Qualité des contextes éducatifs de la petite enfance” research team, at the Center for Research on Intervention and School Success (CRIRES) and in the PÉRISCOPE network. She can be contacted at email: caroline.bouchard@fse.ulaval.ca.







Lorie-Marlène Brault-Foisy     is a full professor in early childhood education didactics at Université du Québec à Montréal (Québec, Canada). She is a member of the “Qualité des contextes éducatifs de la petite enfance” research team. Her current research interest includes early childhood education quality, neuroeducation, and child development. Her publication topics include neuroeducation and executives’ functions. She can be contacted at email: brault-foisy.lorie-marlene@uqam.ca.



Nathalie Bigras     is a full professor in early childhood education didactics at Université du Québec à Montréal (Québec, Canada). She is the scientific director of the “Qualité des contextes éducatifs de la petite enfance” research team. Her work focuses on the early childhood education quality and support for adults who work with young children. Her recent publication topics include the wellbeing of early childhood center’s directors and educators during the COVID-19 pandemic. She can be contacted at email: bigras.nathalie@uqam.ca.



Gilles Cantin     is a retired professor in early childhood education didactics at Université du Québec à Montréal (Québec, Canada). His interests were the quality of early childhood educational settings, the educators-parents partnership, and educators’ practices. He led the research team that contributed to the adaptation of the CLASS tools into French for Quebec and to the development of the Ministry of the Family Assessment and improvement of educational quality of educational childcare service facilities measure. He can be contacted at email: Cantin.gilles@uqam.ca.