

## **Évaluer la compréhension en lecture d'un récit et d'un texte informatif auprès d'élèves de 8 ans**

Catherine Turcotte  
*Université du Québec à Montréal*

Nathalie Prévost  
*Université du Québec à Montréal*

Pier-Olivier Caron  
*Université TÉLUQ*

### **Résumé**

Dans cette étude, des tests de compréhension en lecture d'un récit et d'un texte informatif ont été créés à partir de modèles théoriques communs et en suivant des étapes de conception précises (DeVellis, 2003). Ils furent administrés auprès de 401 élèves francophones de 8 ans. Les données recueillies ont subi des analyses par items, factorielles et multiniveaux. Ceci a permis d'examiner les taux de réussite aux tests, de mieux déterminer leurs composantes et de les comparer afin de comprendre leur portée pour la communauté praticienne et scientifique. Les résultats révèlent que les deux tests ont une structure factorielle semblable. Or, les questions ouvertes suscitant des inférences anaphoriques, causales et lexicales sont plus ardues à réussir, surtout dans le test utilisant un texte informatif. Ces nouveaux instruments peuvent contribuer à dégager des pistes d'enseignement et favoriser la transition entre le premier et le deuxième cycle du primaire en lecture.

### **Abstract**

In this study, reading comprehension tests of narrative and informational text were created using common theoretical models and by following specific design steps (DeVellis, 2003). They were administered to 401 French-speaking 8-year-old students. The data collected underwent item, factorial and multilevel analyses. It led to examine the success rates for each test, to better determine their components and to compare them, in order to understand their usefulness for teachers and researchers. Results reveal that the two tests have a similar factorial structure. However, open-ended questions targeting anaphoric, causal and lexical inferences are more difficult to achieve, especially in the test using an informational text. These new instruments can help identify teaching avenues and promote a better transition between the first and second cycles of primary school in reading.

## Évaluer la compréhension en lecture d'un récit et d'un texte informatif auprès d'élèves de 8 ans

### Introduction et énoncé du problème

Les élèves qui développent des connaissances et des habiletés fondamentales à la lecture, comme la connaissance des lettres, la conscience phonologique, le vocabulaire et les relations lettre-son (Sprenger-Charolles et coll., 2018; Storch et Whitehurst, 2002; Suggate et coll., 2018) sont plus susceptibles de devenir de bons lecteurs. Or, ces derniers ne garantissent pas la mise en œuvre des habiletés de haut niveau nécessaires pour comprendre des textes ou encore, apprendre à partir de ces textes en milieux scolaires ou professionnels (Connor et coll., 2015). D'ailleurs, depuis quelques années, un consensus scientifique émerge à l'égard de la complexité de la compréhension en lecture et du besoin de cerner les divers niveaux de compréhension lors de l'enseignement et de l'évaluation (Catts et Kamhi, 2017).

Cette complexité ne se reflète pas nécessairement dans les instruments d'évaluation de la compréhension en lecture utilisés en classe ou même dans les travaux de recherche. Par exemple, ils évaluent souvent une quantité très limitée d'habiletés, surtout lorsqu'ils sont destinés aux jeunes lecteurs du primaire ayant entre 6 et 9 ans (Sabatini et coll., 2016). Des chercheurs (Keenan et coll., 2008) ont aussi révélé que des évaluations standardisées de la compréhension en lecture couramment utilisées dans la pratique comme dans les recherches mesurent des habiletés de lecture de mots davantage que celles plus directement liées à la compréhension. Il est alors difficile de documenter les progrès des élèves ou encore d'établir les relations possibles entre diverses habiletés liées à la compréhension en lecture.

Pour remédier cette situation, Clemens et Fuchs (2021) soutiennent qu'il devient nécessaire de développer des tests ayant une bonne validité théorique tout en détenant le potentiel d'inspirer les interventions en classe. Ce type d'instrument fournit aux enseignants des informations détaillées sur le cheminement des élèves et sur la mobilisation des connaissances, des habiletés et des stratégies déjà enseignées (Afflerbach et coll., 2018). En d'autres mots, les tests de compréhension en lecture devraient contribuer à répondre à des questions telles : les élèves se sont-ils approprié les objets enseignés, quel enseignement devrait-on planifier pour qu'ils y parviennent ou quelle évaluation supplémentaire devrait-elle être menée pour mieux cerner leurs difficultés?

Or, la conception de tels tests rencontre un second enjeu, cette fois lié au genre de texte (narratif, informatif, argumentatif, etc.). Cet aspect n'est pas considéré dans la majorité des instruments, ce qui mène à des généralisations discutables sur le plan des habiletés des élèves puisque les genres de textes sollicitent des composantes distinctes de la compréhension (Kim et Petscher, 2021). À titre d'exemple, le récit raconte une histoire à l'aide de faits qui s'enchaînent ou en présentant plusieurs épisodes dans la vie d'un personnage, tandis que le texte informatif repose sur la transmission d'informations et de connaissances propres à un domaine (l'astronomie, les reptiles, l'histoire du Québec, etc.). La compréhension du récit s'avère plus influencée par les habiletés de décodage alors que la compréhension des textes informatifs, par les connaissances sur le monde (Best et coll., 2008). En outre, la formulation d'inférences est plus facile dans les récits, qui sont plus familiers, que dans les textes informatifs (Clinton et coll., 2020). De plus, les récits ont tendance à contenir une proportion plus élevée de mots fréquents et connus pour les jeunes

élèves, alors que les textes informatifs utilisent davantage un lexique spécialisé lié à une discipline (histoire, science, etc.).

Ainsi, le récit est plus facilement compris par les jeunes lecteurs. Ceci s'attache peut-être aussi au fait qu'en milieu scolaire, les enseignants sont portés à aborder le récit plus que tout autre genre (Duke et Roberts, 2010). D'ailleurs, même lorsque la compréhension d'un texte oralisé est évaluée, les élèves de la 2<sup>e</sup> année du primaire montrent de plus grandes habiletés avec le récit (Kim et Petschner, 2021). En ce sens, la réussite à un test de compréhension utilisant un récit ne permet pas de présumer des habiletés en compréhension en lecture d'autres genres de textes.

En somme, les milieux scolaires et scientifiques devraient disposer d'évaluations en compréhension en lecture de plus d'un genre de textes et qui reposent sur des modèles théoriques récents. Au Québec, peu d'instruments de ce type existent pour évaluer les habiletés en compréhension en lecture des élèves d'environ 8 ans, en grand groupe ou de façon individuelle. Afin de planifier des séquences d'enseignement ciblant des composantes de la compréhension en lecture ou identifier des élèves qui bénéficieraient d'un soutien ou d'une évaluation supplémentaire, les praticiens ont donc peu d'instruments sur lesquels compter. De plus, à notre connaissance, aucun ne propose une évaluation à partir des deux genres de textes les plus fréquemment rencontrés, c'est-à-dire le récit et le texte informatif, de manière à offrir un portrait plus complet des habiletés en compréhension des jeunes lecteurs. Enfin, le passage du premier cycle vers le deuxième cycle du primaire est marqué au Québec par des exigences nouvelles sur le plan de la compréhension d'éléments implicites et par l'introduction d'une plus grande diversité de textes (MELS, 2011). À l'égard de ces élèves, il serait intéressant de détenir des données sur leur appropriation de connaissances et d'habiletés afin de susciter une transition réussie intercycle.

Pour y parvenir, ces tests devraient détenir une fonction d'amélioration (De Ketele, 2010), c'est-à-dire contribuer à déterminer ce qui est acquis ou demeure à développer afin d'orienter l'enseignement. Ces tests peuvent aussi être employés dans le cadre d'un dépistage universel, en ce sens qu'ils peuvent aider à déterminer si des élèves ont besoin d'une évaluation plus approfondie afin de valider une hypothèse de retard ou de difficultés persistantes (Compton et coll., 2006). Enfin, sur le plan de la passation, les tests administrés en grands groupes sont moins coûteux en temps et en ressources et peuvent ainsi être particulièrement appréciés dans un contexte de pénurie de personnel dans les écoles (Adlof et coll., 2017), qui touche le milieu scolaire au Québec depuis plusieurs années. Les tests développés dans cette étude considéreront tous ces enjeux.

## Objectifs

Cet article a donc comme objectifs a) de décrire le développement de courts tests en compréhension de lecture d'un récit et d'un texte informatif pour des élèves de 7 et 8 ans en s'appuyant sur des modèles théoriques, b) d'examiner les taux de réussite à ces tests afin de les comparer et comprendre leur portée pour les enseignants ou les chercheurs.

## Cadre de référence sur la compréhension en lecture

Des modèles théoriques (Irwin, 2007; Kintsch et Rawson, 2005; Perfetti et Stafura, 2014) représentent la compréhension en lecture telle une interaction entre plusieurs

composantes, allant de la lecture de mots, vers la compréhension de phrases et celle de paragraphes ou de textes entiers. En utilisant les idées et les informations du texte à ces divers niveaux ainsi qu'en formulant des inférences nécessaires pour combler l'implicite, les lecteurs établissent un modèle de situation, c'est-à-dire une représentation mentale du texte. En somme, la compréhension se construit tout au long de la lecture à l'aide de l'intégration progressive d'éléments explicites et implicites et en sollicitant les connaissances de la personne qui lit, ce qui alimente en retour ses connaissances et expériences personnelles.

Selon ces modèles, le lecteur doit lire et comprendre les mots inclus dans les phrases. Ces connaissances et habiletés en lecture de mots et de groupes de mots impliquent aussi la connaissance du lexique employé. Ces connaissances sont nécessaires pour lire et comprendre la base du texte, ce qui est dévoilé par l'auteur. Ceci assure du moins une compréhension de bas niveau, qui se manifeste par des habiletés de repérage textuel. Cette compréhension appuiera les processus de compréhension de haut niveau qui s'ensuivent (Potocki et coll., 2014).

Dans un récit comme dans un texte informatif, le lecteur doit aussi tracer des liens entre les phrases et les idées qu'elles contiennent pour les intégrer et en construire une représentation locale cohérente. Ces processus soutiennent la compréhension de la microstructure du texte. Le lecteur mobilise alors ses connaissances relatives aux marqueurs de relation employés par l'auteur pour exprimer et unir les idées, mais doit aussi faire preuve d'habiletés inférentielles puisque la plupart des textes comprennent une part d'implicite (Elleman, 2017). Afin de maintenir la cohésion des idées du texte, il est donc nécessaire de formuler des inférences anaphoriques, lexicales et causales (Bianco et coll., 2002).

Les inférences anaphoriques consistent à lier un mot de substitution à son référent. Par exemple, le référent "Chaperon Rouge" pourrait être remplacé par divers pronoms, synonymes ou mots génériques (elle, la fillette, cette imprudente, etc.), lesquels doivent être compris par le lecteur. Les inférences lexicales consistent à saisir au moins partiellement le sens d'un mot rare ou inconnu dans un texte en s'appuyant sur des indices. Par exemple, un lecteur pourrait inférer qu'un « fossa » fait partie de la famille des félins puisque l'auteur compare cet animal à un tigre, un lynx et un puma. Un lecteur peut aussi utiliser des indices présents dans un mot pour y attribuer un sens (Fejzo, 2021; Levesque et coll., 2019). Cette utilisation de la morphologie permettrait de comprendre le mot « grandiose » à partir du mot « grand ». Les inférences causales, quant à elles, amènent le lecteur à comprendre les liens entre les causes et les effets d'un phénomène. Cette causalité se retrouve abondamment dans les textes en sciences naturelles et humaines. Or, les auteurs ne précisent pas toujours les marqueurs de relation de causalité (parce que, donc) ou n'utilisent pas toujours des verbes ou noms qui évoquent cette causalité (provoque, crée, suscite, problème, cause, effet, conséquence, etc.). C'est au lecteur de saisir ces liens entre les causes et les effets laissés implicites par l'auteur. Par exemple, dans l'extrait suivant, un lecteur doit inférer que l'école sera fermée à cause d'une tempête et que l'examen de science sera donc reporté.

*Il était tombé plus de 35 centimètres de neige depuis hier soir et des vents allant jusqu'à 70 km/h soufflaient toujours sur la ville. Les enfants réalisèrent avec ravissement que leur examen de sciences serait reporté.*

Ces inférences sont considérées comme des habiletés de haut niveau puisqu'elles interpellent la compréhension au-delà de ce qui est littéralement dévoilé. Elles appuient également la compréhension de la macrostructure (Kintsch et Rawson, 2005), laquelle convoque les macroprocessus de compréhension (Irwin, 2007). Ces derniers touchent l'identification de la structure du texte ainsi que la compréhension des aspects ou des idées principales abordés dans les paragraphes ou les différentes parties du texte. Les macroprocessus exigent donc la coordination de plusieurs connaissances et habiletés. Tout comme les inférences, ils sont donc plus complexes à maîtriser (Landi et Ryherd, 2017; Oakhill et Cain, 2012) et sont aussi qualifiés d'habiletés de « haut niveau » (Block et coll., 2009 ; Wharton-McDonald et Swiger, 2009). Il s'agit également du niveau de compréhension exigeant le plus de connaissances sur le genre de texte employé puisque ce dernier détermine la façon dont les idées s'enchaînent ou les informations s'organisent.

Dans ces mêmes modèles (Irwin, 2007; Kintsch et Rawson, 2005; Perfetti et Stafura, 2014), un autre type d'inférence vient jouer un rôle sur le plan de l'appropriation du texte par tout lecteur. Lorsque ce dernier réfère à une connaissance personnelle ou une expérience vécue pour interpréter une partie du texte, il fait une inférence élaborative. Ce serait le cas dans un récit lorsqu'un lecteur comprend les sentiments de déception et d'amertume vécus par un personnage en se remémorant une expérience personnelle semblable. Dans un texte informatif, ceci peut être observé lorsqu'un lecteur utilise ses connaissances pour comprendre qu'un animal a une importante longévité puisqu'il se situe au sommet de la chaîne alimentaire, même si l'auteur ne mentionne jamais ce fait. En somme, lorsque le lecteur effectue une inférence élaborative, ceci peut indiquer qu'il arrive à créer une représentation du texte lui étant personnelle en mettant à profit ses connaissances.

Tous ces processus peuvent faire l'objet d'un contrôle du lecteur lorsqu'il relève ses bris de compréhension, y remédie ou réfléchit à sa propre compréhension. Ce processus métacognitif est souvent évalué dans un questionnaire autorapporté spécifique à cet effet ou lors d'un entretien individuel centré sur le « comment » un lecteur s'y prend lorsqu'il se prépare à lire ou lorsqu'il fait face à un bris de compréhension, par exemple (voir Mokhtari et Reichard, 2002). Ce processus ne sera pas évalué spécifiquement dans les tests développés au cours de cette étude puisque ceci exige une tout autre forme de questions ou d'interactions avec la personne évaluatrice.

## Méthode

Afin de répondre à l'objectif de développer de courts tests en compréhension de lecture s'adressant à des élèves du début du primaire, donc ayant entre 7 et 8 ans (fin 2<sup>e</sup> et début 3<sup>e</sup> années du primaire), puis d'examiner les taux de réussite à ces tests en fonction du genre de texte, une méthode s'appuyant sur les étapes de conception d'un test a été employée.

## Participants

Puisque cette cueillette s'est déroulée en début d'année scolaire, des élèves de 3<sup>e</sup> année ont été recrutés. En octobre 2020, 401 élèves ont participé, après avoir reçu le certificat éthique de l'université d'attache des personnes responsables du projet ainsi que les formulaires de consentement signés par les parents. Les élèves participants ont un âge

moyen de 8 ans et 9 mois et sont francophones. Ils proviennent de deux centres de services scolaires dans la grande région de Montréal et sont répartis dans 8 écoles de quartiers ayant un niveau socioéconomique considéré comme étant moyen. Il y a légèrement plus de filles que de garçons dans la population étudiée (52% et 48%).

Afin de contrôler la réactivité à la mesure (l'effet d'entraînement) et augmenter la validité interne des tests, les 401 élèves ont complété au moins un test parmi les deux. Parmi ceux-ci, 85 ont fait les deux tests, 255 élèves n'ont fait que le test informatif et 231 n'ont fait que le test récit. Les classes ayant reçu l'un ou l'autre ou les deux tests ont été attribuées aléatoirement. De plus, une procédure de contre-balancement a été utilisée afin de modifier l'ordre de passation : la moitié des élèves ayant été soumis aux deux tests ont d'abord fait celui du récit et ensuite le texte informatif et vice-versa pour l'autre moitié.

### Étapes de conception des tests

Les huit étapes de conception d'un test proposées par DeVellis (2003) ont été respectées puisque cette démarche a pour point de départ l'ancrage à des modèles théoriques, ce qui est cohérent avec les besoins évoqués précédemment en matière de validité théorique des instruments. En ce sens, tout au long de la démarche et à la lumière des performances des élèves à chaque cueillette de données, les items et les textes ont été révisés dans leur forme et dans leur contenu. Les étapes de conception ainsi que le calendrier sont présentés dans le tableau 1. Les données recueillies dans la présente étude sont donc issues de la huitième et dernière étape du calendrier.

**Tableau 1**

*Étapes de conception des tests selon le modèle de DeVellis (2003)*

Étapes de DeVellis (2003)	Procédure	Calendrier
1) déterminer clairement ce qui est à mesurer en fonction d'un modèle théorique	Sélection du modèle théorique et des textes à faire lire dans les tests	Septembre et octobre 2018
2) générer des items et	Création d'items à partir du modèle théorique et selon les	Novembre 2018
3) déterminer le type d'items	deux genres de textes	
4) vérifier la clarté et la pertinence des items à l'aide d'expertes	Consultation d'une chercheuse externe et d'une orthopédagogue	Janvier 2019
5) ajuster les items	Révision des items en fonction des commentaires des expertes	Février et mars 2019
6) soumettre les items à un groupe représentant la population cible	Passation auprès de 238 élèves	Mai 2019
7) analyser les items et la validité de construit	Analyse des réponses, modification du texte et des items	Septembre 2019 à février 2020
8) améliorer la fidélité du test.	Passation de la version révisée auprès de 401 élèves	Octobre 2020

## Forme et contenu des tests

Le contenu des tests est inspiré des modèles théoriques présentés plus tôt dans cet article. Chaque test comprend donc des questions de repérage d'informations, d'inférences anaphoriques, lexicales, causales et élaboratives puis des questions sur la macrostructure du texte propre à chaque genre : récit et informatif. En ce qui concerne la forme des items, il faut rappeler que le test s'adresse à des élèves d'environ 8 ans (fin de 2<sup>e</sup> année ou début de 3<sup>e</sup> année du primaire). Leur apprentissage de toutes les composantes de l'écriture, dont la syntaxe, l'orthographe, la grammaire, la ponctuation, n'est pas maîtrisé. Pour ne pas surcharger leur activité en exigeant seulement des réponses à des questions ouvertes à l'écrit, des questions ayant des choix multiples ont été insérées en alternance avec des questions ouvertes.

Un récit intitulé « Le malin petit cochon » a été écrit et révisé de manière à présenter une courte histoire qui suscite différentes inférences à partir de la structure suivante : une situation initiale, un élément déclencheur, des péripéties et un dénouement menant à une situation finale. Le lettrage est régulier et script. La version finale contient 15 phrases et 225 mots, dont 116 mots différents. Ce texte est divisé en 4 courts paragraphes correspondant aux étapes du récit. Les phrases possèdent en moyenne 13 mots et peuvent donc faire plus d'une ligne. Le vocabulaire emploie plusieurs mots connus et de haute fréquence (loup, cochon, il, petit, pour, que qui, son, un, plume, bien, etc.) tels que mentionnés dans la liste orthographique (MEES, 2019), liste publiée en complément du programme de formation de l'école québécoise. Le texte introduit également des mots et expressions littéraires (astucieux, stratagème, faim de loup), en plus de quelques mots polysyllabiques (diaboliques, discrètement). L'histoire contient des péripéties reliées à une intrigue unique. Les verbes sont à l'imparfait et au passé simple, ce qui correspond à l'emploi des temps de verbe dans la structure traditionnelle d'un récit. Ces critères sont analogues à ceux présentés dans l'échelle de complexité des textes de Fountas et Pinnell (1996) au niveau L, lequel s'adresse aux élèves des 2<sup>e</sup> et 3<sup>e</sup> années du primaire.

Les items suivants sont insérés dans le questionnaire sur le texte :

- Question 1 : question à choix multiple faisant appel aux habiletés de repérage
- Question 2 : question ouverte faisant appel aux habiletés de repérage
- Question 3 : question à choix multiple faisant appel aux inférences anaphoriques
- Question 4 : question ouverte faisant appel aux inférences anaphoriques
- Question 5 : question à choix multiple faisant appel aux inférences causales
- Question 6 : question ouverte faisant appel aux inférences causales
- Question 7 : question à choix multiple faisant appel aux inférences lexicales
- Question 8 : question ouverte faisant appel aux inférences lexicales
- Question 9 : question ouverte faisant appel à une inférence élaborative
- Question 10 à 12 : questions sur le schéma du récit (personnages, problème, péripétie)
- Question 13 : question de remise en ordre d'étapes du récit (macrostructure)
- Question 14 : question à choix multiples sur le genre de texte (macrostructure)

Le test de compréhension d'un texte informatif est lié au texte « Pourquoi la neige tombe-t-elle ? » (Perreault, 2017). Il s'agit d'un court texte comportant un seul paragraphe trouvé dans le magazine scientifique québécois destiné aux 6-9 ans « Les Explorateurs ». Il aborde la formation de la neige et comprend 119 mots, dont 73 différents. Ces derniers sont

réunis dans 7 phrases, lesquelles possèdent donc une moyenne de 14 mots. Le lettrage est le même que dans le récit. Le vocabulaire emploie plusieurs mots connus et de haute fréquence (neige, nuage, étoile, hiver, plus, eau, sont, des, ces, etc.) selon la même liste orthographique susmentionnée (MEES, 2019), mais introduit quelques mots polysyllabiques moins courants spécifiques au phénomène décrit (géométrique, température, humidité). Les verbes sont essentiellement au présent, sauf un verbe à l'impératif (amuse-toi) et un second au futur simple (tu verras). Ces temps de verbes sont aussi ceux utilisés dans les textes informatifs présentés aux élèves ayant l'âge ciblé. Ces critères correspondent aussi au niveau L de l'échelle de Fountas et Pinnell (1996).

Le test contient moins de questions que celle du récit puisque le récit exige des questions sur les différentes étapes du schéma de récit alors que ceci ne concorde pas avec le texte informatif. Les 11 questions sont donc présentées comme suit :

- Question 1 : question à choix multiple faisant appel aux habiletés de repérage
- Question 2 : question ouverte faisant appel aux habiletés de repérage
- Question 3 : question à choix multiple faisant appel aux inférences anaphoriques
- Question 4 : question ouverte faisant appel aux inférences anaphoriques
- Question 5 : question à choix multiple faisant appel aux inférences causales
- Question 6 : question ouverte faisant appel aux inférences causales
- Question 7 : question à choix multiple faisant appel aux inférences lexicales
- Question 8 : question ouverte faisant appel aux inférences lexicales
- Question 9 : question ouverte faisant appel à une inférence élaborative
- Question 10 : aucune
- Question 11 : aucune
- Question 12 : aucune
- Question 13 : question à choix multiples sur le sujet du texte (macrostructure)
- Question 14 : question à choix multiples sur le genre de texte (macrostructure)

Les tests furent corrigés par trois correctrices. La première détenait une maîtrise et les deux autres, un doctorat en sciences de l'éducation. Elles possédaient aussi un brevet d'enseignement au primaire du Québec. En ce qui a trait aux questions ouvertes, les correctrices ont d'abord listé les réponses adéquates et erronées des élèves pour chaque item à partir de 30 tests complétés pour chaque texte. Cela leur a permis d'échanger sur les raisons justifiant l'attribution ou non d'un score de réussite, en fonction des modèles théoriques présentés plus haut dans cet article. Une fois ces précisions apportées, les évaluatrices ont corrigé de façon indépendante. Tout au long de cette correction, lorsque des réponses inédites apparaissaient, les trois correctrices s'entendaient à nouveau sur le score à attribuer. Cette procédure a permis aux trois correctrices d'avoir exactement le même référentiel de réponses acceptables. Les textes, les questions et les corrigés issus de ce travail sont gratuitement disponibles en ligne à l'adresse suivante :

<https://adel.uqam.ca/docs-category/comprehension-de-recits-et-de-textes-informatifs/>

Enfin, pour des fins de clarté dans cet article et puisqu'il y a des questions spécifiques au récit, les items 10, 11 et 12 ne comprennent pas de données pour le test utilisant un texte informatif.

## Résultats

Les résultats présentent d'abord des statistiques descriptives ainsi que les données d'une analyse par items et par processus. Ensuite, les résultats d'une analyse factorielle et d'une analyse multiniveau sont présentés.

### Analyses descriptives

Afin d'obtenir une première analyse globale, le tableau 2 présente les moyennes à chaque test. Lorsque les items réservés uniquement au récit sont retirés de l'analyse (voir Malin<sup>a</sup>) la moyenne de réussite demeure néanmoins plus élevée pour ce test.

**Tableau 2**

*Statistiques descriptives des deux tests de compréhension en lecture*

Statistique	Malin	Malin <sup>a</sup>	Neige	Diff
<i>n</i>	231	231	255	85
Étendue <sup>b</sup>	0-14	0-11	0-11	
Moyenne	10.20	7.87	7.47	-.682
Écart type	2.71	2.22	2.14	2.00
Erreur standard	.178	.146	.134	.216
Réussite (en %)	72.86	71.55	67.91	6.20

*Note.*

<sup>a</sup> Version de Malin équivalente à Neige (11 items analogues);

<sup>b</sup> Étendue théorique

### Analyse par items

Afin d'examiner si les items des deux tests ont des niveaux de difficulté semblables, une première analyse par item est effectuée, selon le genre de texte. Le tableau 3 présente ces données. Bien que deux items puissent évaluer un même processus, il est important de comprendre d'abord le taux de réussite par item, étant donné que leur format (ouvert ou à choix multiples) diffère. Par exemple, le traitement de l'inférence anaphorique est évalué à partir des items 3 et 4, mais l'item 3 est à choix multiples alors que l'item 4 est une question ouverte.

**Tableau 3**

*Pourcentage de réussite aux items pour les deux tests (en %)*

Test	Item													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Récit (malin)	93	81	66	52	86	46	69	72	73	93	60	80	80	70
Informatif (neige)	88	94	73	39	55	52	95	39	73				65	74

En premier lieu, les items 1, 2, 9 et 14 semblent moins ardues que les autres items, avec un taux de réussite de 70% et plus, peu importe le genre de texte à lire (récit ou informatif). Les items 1 et 2 examinent les capacités de repérage des élèves, tandis que l'item 9 sollicite la formulation d'une inférence élaborative. Quant à l'item 14, réussi par

une majorité de participants, il s'agit d'une question visant la compréhension de l'intention de l'auteur et du genre de texte. Les items 1 et 14 sont à choix multiples alors que les items 2 et 9 sont des questions ouvertes.

Les items 4 et 6 semblent, quant à eux, plus ardues également dans les deux tests, avec des taux de réussite ne dépassant pas 52%. Ces deux items sont des questions faisant appel aux inférences anaphoriques et causales. Il s'agit de deux questions ouvertes.

En deuxième lieu, les pourcentages de réussite aux items indiquent certaines différences entre les deux tests. Des taux de réussite plus faibles sont particulièrement observés pour deux items dans le test utilisant le récit (malin) pour les items 2 (question ouverte de repérage) et 7 (choix multiples inférence lexicale), avec des pourcentages de réussite plus faibles de 13% et 26%.

Or, des taux de réussite sont substantiellement plus faibles pour un nombre plus élevé d'items venant du test utilisant le texte informatif (neige). Il s'agit des items: 4, 5, 8 et 13, avec des taux de réussite plus faibles respectivement de 13%, 31%, 33% et 15%. L'item 5, une question à choix multiple suscitant une inférence causale, est réussi d'ailleurs par 55% des élèves, contrairement à 86% pour le test sur le récit. L'item 8, une question ouverte suscitant la formulation d'une inférence lexicale, est d'ailleurs moins bien réussi spécifiquement dans le test utilisant le texte informatif (neige) avec 39% de réussite. Les trois items présentant une question ouverte (4, 6 et 8) requérant les trois types d'inférences de cohésion (anaphorique, causale, lexicale) s'avèrent donc les moins bien réussis, surtout pour le test utilisant un texte informatif. En ce qui concerne l'item 13 du test ayant un texte informatif (neige), le taux de réussite de 65% présente une différence de 15% avec le taux de réussite du test récit (malin).

### Analyse par processus

Cette analyse regroupe les items par processus. Le tableau 4 donne à voir que les pourcentages de réussite sont plus élevés pour le récit en ce qui concerne quatre processus : les inférences anaphoriques, causales et lexicales, puis pour la macrostructure.

**Tableau 4**

*Pourcentage de réussite par processus évalué (en %)*

Processus évalués	Items	Récit (malin)	Texte informatif (neige)
Repérage	1,2	87	91
Inférence anaphorique	3,4	59	56
Inférence causale	5,6	66	53,5
Inférence lexicale	7,8	70,5	67
Inférence élaborative	9	73	73
Macrostructure	13,14	75%	69,5

### Analyse factorielle

Une analyse de la structure factorielle par l'intermédiaire d'une analyse en composantes principales des questionnaires liés aux tests fut aussi réalisée, avec l'hypothèse qu'il n'y aura qu'une seule dimension à retenir par test. L'analyse parallèle suggère en effet la présence d'un seul facteur pour le test de compréhension du récit (malin) et celui du texte informatif (neige), qui explique 22.94% et 21.25% de la variance

respectivement. Le tableau 5 montre la structure factorielle obtenue et leur cohérence interne évaluée par le  $\alpha$  de Cronbach.

**Tableau 5**

*Structures factorielles des deux tests de compréhension en lecture*

Item	Récit (malin)	Informatif (neige)
1	.434	.428
2	.559	.184
3	.496	.488
4	.521	.611
5	.563	.397
6	.493	.447
7	.158	.123
8	.404	.436
9	.578	.580
13	.446	.474
14	.473	.621
$\alpha$ de Cronbach	.640	.618

Le tableau 5 indique que les structures factorielles des deux tests sont adéquates. En général, on note une structure factorielle très similaire entre les deux tests, ce qui est souhaitable puisque les tests sont issus d'un modèle théorique commun. Les  $\alpha$  de Cronbach de .640 et de .618 peuvent sembler peu élevées, mais il est utile de rappeler à cet effet que les tests sont composés d'items sollicitant des processus distincts et possédant des niveaux de complexité différents.

Par ailleurs, les données du tableau 5 indiquent que l'item 7 aux deux tests ne se lie pas très bien avec les autres items. Il s'agit d'une question à choix multiples exigeant une inférence lexicale. Or, le retrait de l'item 7 n'améliore que peu l' $\alpha$ , soit .657 pour le test utilisant un récit (malin) et .622 pour le test utilisant un texte informatif (neige). L'item 2 dans le test sur le texte informatif (neige) ne se lie pas non plus très bien avec les autres items. Il s'agit d'une question de repérage ayant cette fois une forme ouverte.

### Analyse multiniveau

Puisque l'étude comporte des composantes à traiter simultanément, soit intra participant (participants ayant réalisé les deux tests), inter participant (participants ayant réalisé un seul test), en plus de provenir de classes (élèves d'une même classe), l'analyse multiniveau a été choisie (Duplessis-Marcotte et coll., 2022). Il faut noter ici que les tests sont considérés a priori comme des variables analogues, ce qui justifie les présentes analyses. Par simplicité et afin de mieux refléter les résultats, les résultats totaux (sur 11) des deux tests sont utilisés.

Trois modèles sont envisagés pour être comparés. Le premier vise à tester l'effet des participants sur leur score total. L'ordonnée est dans ce modèle un effet aléatoire qui représente le score attendu de l'élève. C'est cet effet qui pondère la variance intra participant, dans le cas où le participant a réalisé les deux tests. Le deuxième modèle teste l'ajout de la classe, soit l'effet d'appartenir à un groupe-classe. Ces deux premiers modèles

servent à tenir compte (contrôler) de l'effet de intrasujet et de la classe pour mieux observer l'effet du genre de texte. Le troisième effet consiste à évaluer l'effet du genre de texte, soit le récit (malin) relativement au texte informatif (neige). Il s'agit d'un effet fixe ajouté au modèle précédent qui représente l'effet du test sur le score des élèves. Les résultats sont présentés dans le tableau 6.

**Tableau 6***Résultats des analyses multiniveaux*

	Modèle 1	Modèle 2	Modèle 2
AIC	2116.6	2114.1	2105.2
BIC	2129.2	2130.9	2126.2
Pseudo R2	53.23	54.39	59.39
Nombre de paramètres	3	4	5
Effet fixe (ordonnée)	7.663 (.106)*	7.695 (.142)*	8.025 (.177)*
Effet fixe (histoire)			-.620 (.183) *

Note. \*=  $p < .001$

Pour ces analyses, il y a 486 observations, dont 401 participants (groupement de niveau 2). Il y a 316 élèves avec une donnée (un score de test) et 85 avec deux données (deux scores de tests). La progression des modèles montre que a) le modèle 2 est supérieur au premier,  $\chi^2(1) = 4.479, p = .034$ , ce qui souligne un effet significatif du groupe-classe, et que b) le modèle 3 est significativement supérieur au modèle 2,  $\chi^2(1) = 10.881, p < .001$ , ce qui soutient l'effet relatif au genre de texte. Comme attendu, les critères d'informations AIC et BIC diminuent avec l'augmentation du nombre de paramètres. La pseudovariance expliquée, quant à elle, est à sa plus élevée au modèle 3, soit 59.39%, une augmentation significative de 5% par rapport au modèle 2.

Ces résultats montrent la pertinence d'ajouter le genre de texte comme variable au modèle. Le test a un effet sur le rendement des élèves et celui qui comporte un texte informatif à lire s'avère plus difficile. En effet, le modèle 3 montre que les élèves ont de moins bons résultats au test de compréhension utilisant un texte informatif (neige) qu'au test utilisant un récit (malin) (cette dernière étant le test de référence),  $\beta = -.620, p < .001$ .

## Discussion

Cette étude avait pour objectif de créer de courts tests de compréhension s'appuyant sur des modèles théoriques (Irwin, 2007; Kintsch et Rawson, 2005) et destinés aux élèves des 2<sup>e</sup> et 3<sup>e</sup> années du primaire (environ d'âge 8 ans). Deux tests ont été créés: le premier porte sur la compréhension d'un récit et l'autre porte sur un texte informatif.

Dès le départ, les tests ont été créés en fonction de modèles théoriques afin de pouvoir arrimer l'évaluation à l'enseignement de composantes précises ou encore identifier des élèves pour lesquels une évaluation plus approfondie de ces composantes serait nécessaire. Les deux tests ont également été analysés par deux expertes afin d'effectuer une première critique de leur validité théorique. Les versions ont été alors modifiées grâce à

leurs commentaires, puis une seconde fois après une passation à une première cohorte d'élèves de 7 et 8 ans. Toutes ces étapes, qui sont conformes au modèle de conception des tests de Devellis (2003), ont permis de s'assurer que chaque item évaluait bien ce qu'il prétendait évaluer en fonction des théories et en fonction de l'âge des élèves.

À la suite d'un dernier tour de passation auprès de 401 élèves de 8 ans, les résultats des analyses factorielles indiquent que les deux tests sont similaires et qu'un seul facteur pour le test de compréhension du récit (malin) et celui du texte informatif (neige) explique 22.94% et 21.25% de la variance. Puisque ces tests ont été conçus à partir de mêmes modèles théoriques, cette similarité était attendue. Certains items se lient moins bien aux autres, mais de les retirer n'améliorerait pas le modèle statistique. De plus, comme ils sont liés à des composantes précises des modèles adoptés, ils peuvent être utiles pour comprendre les conduites d'un jeune lecteur ou dégager des pistes d'intervention ou d'évaluations subséquentes.

### **Difficulté des items selon le genre de texte et la forme des questions**

Sur le plan de la difficulté des items, ceux ayant suscité les plus hauts taux de réussite ciblent la compréhension littérale (items 1 et 2) dans les deux textes. Ce résultat rappelle que ces habiletés dites « de bas niveaux » sont les plus faciles à maîtriser lors des premiers apprentissages de la compréhension en lecture (Potocki et coll., 2014). Malgré le taux élevé de réussite, ces items s'avèrent particulièrement utiles pour identifier des élèves n'ayant toujours pas développé des habiletés de compréhension fondamentales, ce qui peut être tributaire de difficultés à lire ou comprendre les mots ou encore à une lecture peu fluide. Les élèves de 8 ans n'ayant pas répondu adéquatement à ces questions sont donc susceptibles de se situer parmi ceux accusant un retard en lecture et en ce sens, il pourrait être pertinent de poursuivre la démarche d'évaluation afin de cibler des connaissances et des habiletés à développer ou à consolider.

D'autres items se révèlent plus ardues dans les deux tests. En effet, certains items exigeant la formulation d'inférences sont moins bien réussis par les élèves. Ceci rappelle que les habiletés de plus haut niveau sont plus complexes à maîtriser (Landi et Ryherd, 2017; Oakhill et Cain, 2012). Il n'est d'ailleurs pas attendu que les élèves de cet âge réussissent toutes les questions suscitant de telles inférences puisqu'ils seront amenés à y travailler davantage à partir du 2<sup>e</sup> cycle du primaire (MELS, 2011). Rappelons que ce test n'a pas une fonction diagnostique ou certificative, mais permet plutôt de planifier des situations d'enseignement ou des actions à entreprendre pour assurer le développement de connaissances et habiletés soutenant la compréhension en lecture de récits et de textes informatifs.

Il est intéressant d'ailleurs d'observer que les items les plus difficiles sont des questions ouvertes qui exigent la formulation d'inférences anaphoriques, causales ou lexicales et que les taux de réussite sont moindres pour le texte informatif. De plus, l'identification du sujet du texte informatif (voir les items 13) s'avère plus difficile qu'une habileté semblable propre à la macrostructure du récit: la remise en ordre du récit. Cette analyse des habiletés liées à la macrostructure et exigeant la maîtrise du genre de texte montre, en outre, que le genre de texte informatif est moins bien maîtrisé chez ces élèves. En somme, le test utilisant un texte informatif est plus difficile à réussir pour les élèves et ceci rappelle, notamment, qu'il est moins familier (Duke et Roberts, 2010).

Ainsi, le genre du texte utilisé représente une variable pertinente à considérer puisque cette dernière a un effet sur la performance des élèves; le test utilisant un texte informatif s'avérant plus difficile et cette différence est significative. Les taux de réussite dévoilent ainsi une plus grande aisance envers le test utilisant le récit, bien que ce dernier soit plus long afin d'y présenter une histoire ayant une situation initiale, un élément déclencheur, des péripéties et un dénouement menant à une situation finale. Ceci concorde avec plusieurs études ayant révélé que ces deux genres mobilisaient des connaissances et habiletés distinctes, le texte informatif faisant davantage appel aux connaissances sur le monde, aux inférences complexes et au vocabulaire lié aux disciplines (Clinton et coll., 2020; Kim et Petschner, 2021). L'étude présente confirme donc l'importance d'avoir un instrument permettant d'évaluer la compréhension du texte informatif puisque celle se rapportant au récit ne peut être généralisée.

En ce qui a trait à la forme des questions, les élèves semblent détenir une plus grande facilité face aux questions à choix multiples (voir les items 3, 5 et 7). À ce chapitre, il faut noter que les questions d'inférences à choix multiples s'avèrent plus faciles puisque la forme de la question génère l'inférence pour l'élève et ce dernier n'a qu'à la sélectionner parmi d'autres réponses moins adéquates. Dans une question ouverte, l'élève doit formuler lui-même une inférence, en s'appuyant sur des parties du texte qu'il doit mettre lui-même en relation ou encore en saisissant ce que l'auteur ne dévoile pas explicitement (Elleman, 2017). Néanmoins, ceci revêt un certain intérêt du point de vue de l'enseignement puisque les jeunes lecteurs réussissant les items à choix multiples pourraient être considérés comme étant en voie de développer des habiletés lorsque des indices sont offerts.

Il en va de même pour les items liés à la macrostructure. Alors qu'elles sont considérées comme des habiletés de haut niveau à maîtriser, les élèves ont démontré un taux relativement élevé de réussite, surtout pour le récit, aux items 13 et 14. Ces questions étant à choix multiples, une hypothèse sur la forme des items est à soutenir. En effet, pour des élèves de 8 ans, il peut être difficile d'identifier le sujet d'un paragraphe, de rappeler l'ordre d'une histoire de façon complète à l'écrit ou encore de déterminer le genre ou la structure du texte lu à partir de questions ouvertes. Ces dernières requièrent la formulation d'une réponse et sont reconnues pour être plus exigeantes cognitivement, car elles ne fournissent pas d'indices ou d'informations à récupérer en mémoire (Graesser et coll., 2010). De plus, l'écriture d'une réponse, même sous forme de courte phrase, sollicite des connaissances en orthographe et en syntaxe qui ne sont pas maîtrisées par tous au début de deuxième cycle du primaire.

En somme, les questions à choix multiples peuvent être incontournables dans le contexte de ces tests destinés aux jeunes lecteurs de 7 et 8 ans lors d'une passation en grand groupe. Or, l'analyse des forces et des difficultés des élèves ne peut se faire sans tenir compte des indices qu'elles fournissent et qui rendent ces questions plus faciles.

### **Contributions et limites**

Ces tests peuvent être utiles pour les chercheurs et les praticiens qui souhaitent obtenir des informations détaillées sur leurs connaissances ou habiletés des élèves lorsqu'ils comprennent un récit ou un texte informatif (Afflerbach et coll., 2018). Plusieurs auteurs avaient d'ailleurs révélé l'importance de développer des instruments ayant des qualités psychométriques acceptables, en s'assurant d'une bonne validité théorique et la possibilité

de rencontrer les besoins et enjeux du terrain (Clemens et Fuchs, 2021; O'Reilly et coll., 2014). Ces tests font partie de cette nouvelle génération d'instruments.

En effet, en connaissant bien la portée de chaque forme de question et ce que chaque item évalue à partir des modèles théoriques, les praticiens et chercheurs peuvent utiliser ces instruments consciencieusement pour développer des interventions, décider si des élèves nécessitent des évaluations plus approfondies en compréhension ou encore lier les performances à ces tests à celles d'autres types d'évaluations.

Sur le plan statistique, la faible quantité de données par élèves (groupement), plus particulièrement le nombre élevé d'élèves avec une seule observation, pourrait représenter une limite, sachant notamment que de grandes variations peuvent exister sur le plan des habiletés en lecture au sein même d'une seule classe. Or, les analyses montrent qu'il y a réduction de l'erreur standard (voir Tableau 3 comparé à 4) qui provient d'une meilleure estimation de la variance inter participant obtenue à l'aide de l'analyse multiniveau. Celle-ci atteint donc son objectif. Cette préoccupation limite également aussi le nombre de modèles envisageables. Ainsi, d'autres modèles plus complexes que les deux présentés dans le dernier tableau ne sauraient être envisagés.

Enfin, aucun test ne saurait évaluer en profondeur, comme en perspective, l'ensemble des connaissances et habiletés des élèves en compréhension en lecture. En ce sens, ces instruments peuvent être utiles surtout si, d'une part, les réponses des élèves sont analysées en fonction des composantes des modèles théoriques et d'autre part, sont accompagnés d'outils complémentaires au besoin, afin de dresser un profil de jeunes lecteurs et lectrices francophones en plein développement.

Pour aller plus loin, bien que plusieurs items parallèles entre les deux tests possèdent des poids similaires, il serait intéressant de vérifier si les dimensions sont les mêmes entre les deux tests. Aussi, dans une phase subséquente de ce projet, des versions parallèles à ces tests seront créées afin de pouvoir effectuer un suivi entre la 2<sup>e</sup> et la 3<sup>e</sup> année du primaire et mieux observer le progrès de ces jeunes lecteurs entre les cycles scolaires en fonction des deux genres de textes.

## Bibliographie

- Adlof, S. M., Scoggins, J., Brazendale, A., Babb, S. et Petscher, Y. (2017). Identifying children at risk for language impairment or dyslexia with group-administered measures. *Journal of Speech, Language, and Hearing Research*, 60(12), 3507-3522. [https://doi.org/10.1044/2017\\_JSLHR-L-16-0473](https://doi.org/10.1044/2017_JSLHR-L-16-0473)
- Afflerbach, P., Cho, B.-Y., Crassas, M. E. et Kim, J.-Y. (2018). Best practices in reading assessment. Dans L.M. Morrow et L.B. Gambrell (dir.), *Best practices in literacy instruction* (6<sup>e</sup> éd., p. 309-333). Guilford Press.
- Best, R. M., Floyd, R. G. et McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, 29(2), 137-164. <https://doi.org/10.1080/02702710801963951>
- Bianco, M., Coda, M. et Gourgue, D. (2002). *Compréhension*. Éditions de la Cigale.
- Block, C. C., Parris, S. R., Reed, K. L., Whiteley, C. S. et Cleveland, M. D. (2009). Instructional approaches that significantly increase reading comprehension. *Journal of Educational Psychology*, 101(2), 262-281. <https://doi.org/10.1037/a0014883>
- Catts, H. W. et Kamhi, A. G. (2017). Prologue: Reading comprehension is not a single ability. *Language, Speech, and Hearing Services in Schools*, 48(2), 73-76. [https://doi.org/10.1044/2017\\_LSHSS-16-0033](https://doi.org/10.1044/2017_LSHSS-16-0033)
- Clemens, N. H. et Fuchs, D. (2021). Commercially developed tests of reading comprehension: Gold standard or fool's gold? *Reading Research Quarterly*, 57(2), 385-397. <https://doi.org/10.1002/rrq.415>
- Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E. et Seipel, B. (2020). Inferential comprehension differences between narrative and expository texts: a systematic review and meta-analysis. *Reading and Writing* 33(9), 2223-2248. <https://doi.org/10.1007/s11145-020-10044-2>
- Compton, D. L., Fuchs, D., Fuchs, L. S. et Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, 98(2), 394-409. <https://doi.org/10.1037/0022-0663.98.2.394>
- Connor, C. M., Radach, R., Vorstius, C., Day, S. L., McLean, L. et Morrison, F. J. (2015). Individual differences in fifth graders' literacy and academic language predict comprehension monitoring development: an eye-movement study. *Scientific Studies of Reading*, 13(3), 281-300. <https://doi.org/10.1080/10888438.2014.943905>
- DeVellis, R. F. (2003). *Scale development: Theory and applications* (2<sup>e</sup> éd.). SAGE Publications.
- Duke, N. K. et Roberts, K. L. (2010). The genre-specific nature of reading comprehension. Dans D. Wyse, R. Andrews, J. Hoffman (dir.), *The Routledge international handbook of English, language and literacy teaching* (p. 98-110). Routledge.
- Duplessis-Marcotte, F., Lapointe, R. et Caron, P.-O. (2022). Une introduction aux modèles de régressions multiniveaux avec R. *The Quantitative Methods for Psychology*, 18(2), 168-180. <https://doi.org/10.20982/tqmp.18.2.p168>.
- De Ketele, J. M. (2010). Ne pas se tromper d'évaluation. *Revue française de linguistique appliquée*, 15(1), 25-37. <https://doi.org/10.3917/rfla.151.0025>
- Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic

- review. *Journal of Educational Psychology*, 109(6), 761-781.  
<https://psycnet.apa.org/doi/10.1037/edu0000180>
- Fejzo, A. (2021). The contribution of morphological awareness to vocabulary among L1 and L2 French-speaking 4th-graders. *Reading and Writing: An Interdisciplinary Journal*, 34(3), 659–679
- Fountas, I. C. et Pinnell, G. S. (1996). Guided reading: Good first teaching for all children. Heinemann.
- Graesser, A. C., Ozuru, Y. et Sullins, J. (2010). What is a good question? Dans M. McKeown et L. Kucan (dir.), *Bringing reading research to life* (p. 112-141). Guilford Press.
- Irwin, J. W. (2007). *Teaching reading comprehension processes* (3<sup>e</sup> éd.). Prentice-Hall.
- Keenan, J. M., Betjemann, R. S. et Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, 13(3), 281-300.  
<https://doi.org/10.1080/10888430802132279>
- Kim, Y. S. G. et Petscher, Y. (2021). Influences of individual, text, and assessment factors on text/discourse comprehension in oral language (listening comprehension). *Annals of Dyslexia*, 71(2), 218-237. <https://doi.org/10.1007/s11881-020-00208-8>
- Kintsch, W. et Rawson, K. A. (2005). Comprehension. Dans Snowling, M.J. et Hulme, C. (dir.), *Blackwell handbooks of developmental psychology. The science of reading: A handbook*. (p. 209–226). Blackwell Publishing.
- Landi, N. et Ryherd, K. (2017). Understanding specific reading comprehension deficit: A review. *Language and Linguistics Compass*, 11(2), e12234.  
<https://doi.org/10.1111/lnc3.12234>
- Levesque, K. C., Kieffer, M. J. et Deacon, S. H. (2019). Inferring meaning from meaningful parts: The contributions of morphological skills to the development of children's reading comprehension. *Reading Research Quarterly*, 54(1), 63-80.
- Ministère de l'Éducation, du loisir et du Sport (MELS). (2011). *Cadre d'évaluation des apprentissages. Français langue d'enseignement. Enseignement primaire*. Gouvernement du Québec.
- Ministère de l'Éducation et de l'Enseignement supérieur (MEES). (2019). *Liste orthographique. Document de référence*. Gouvernement du Québec.
- Mokhtari, K. et Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, 94(2), 249–259.  
<https://doi.org/10.1037/0022-0663.94.2.249>
- Oakhill, J. V. et Cain, K. E. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, 16(2), 91–121. <https://doi.org/10.1080/10888438.2010.529219>
- O'Reilly, T., Weeks, J., Sabatini, J., Halderman, L. et Steinberg, J. (2014). Designing reading comprehension assessments for reading interventions: How a theoretically motivated assessment can serve as an outcome measure. *Educational Psychology Review*, 26(3), 403-424. <https://doi.org/10.1007/s10648-014-9269-z>
- Perfetti, C. et Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22-37.  
<https://doi.org/10.1080/10888438.2013.827687>
- Perreault, S. (2017). Pourquoi la neige tombe-t-elle ? 7 faits et questions sur la neige. *Les Explorateurs*. Décembre. Bayard éditeur.

- Potocki, A., Bouchafa, H., Magnan, A. et Ecalle, J. (2014). Évaluation de la compréhension écrite de récits chez l'enfant de 7 à 10 ans : vers des profils de compreneurs. *European review of applied psychology*, 64(5), 229-239. <https://doi.org/10.1016/j.erap.2014.08.001>
- Sabatini, J. P., Halderman, L. K., O'Reilly, T. et Weeks, J. P. (2016). Assessing comprehension in kindergarten through third grade. *Topics in Language Disorders*, 36(4), 334-355. <https://doi.org/10.1097/TLD.0000000000000104>
- Sprenger-Charolles, L., Desrochers, A. et Gentaz, E. (2018). Apprendre à lire-écrire en français. *Langue française*, 199(3), 51-67.
- Storch, S. A. et Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, 38(6), 934-947. <https://doi.org/10.1037/0012-1649.38.6.934>
- Suggate, S., Schaughency, E., McAnally, H. et Reese, E. (2018). From infancy to adolescence: The longitudinal links between vocabulary, early literacy skills, oral narrative, and reading comprehension. *Cognitive Development*, 47, 82-95
- Wharton-McDonald, R. et Swiger, S. (2009). Developing higher order comprehension in the middle grades. Dans Israel, S. E. et Duffey, G. G. (dir.) *Handbook of research on reading comprehension* (p. 510- 530). Routledge.