

**Articulation des principes d'évaluation formative dans des guides pédagogiques de
l'Ontario au cycle intermédiaire**

Roméo Fola

Faculté d'éducation

Thèse soumise à la faculté des études supérieures et postdoctorales de l'université d'Ottawa

en vue de l'obtention du grade de

Maitrise ès arts en éducation - M.A. (Éd.)

Comité de thèse :

Marielle Simon, Ph. D., directrice
Éric Dionne, Ph. D., codirecteur
Liliane Dionne, Ph. D., membre
Dany Laveault, Ph. D., membre

Remerciements

Une thèse de maîtrise est un défi. Pour mon cas, relever ce défi aurait à coup sûr été impossible sans l'aide de nombreuses personnes.

Ma gratitude va à l'endroit :

de ma directrice de thèse, Marielle Simon, Ph. D., qui m'a accompagné et soutenu pendant cette grande aventure. Marielle, tu m'as donné le goût de la recherche. À l'aide de techniques constructivistes, tu m'as guidé pas à pas jusqu'au bout. Ta rigueur m'a poussé à toujours donner le meilleur de moi et ta grande patience ainsi que tes encouragements m'ont soutenu durant les phases plus éprouvantes de mon parcours. Les mots ne suffiraient pas pour te témoigner ma reconnaissance. Que le Bon Dieu te le rende au centuple!

de mon codirecteur de thèse, Éric Dionne, Ph. D., pour ses rétroactions enrichissantes. Merci, Éric, pour ton encadrement durant ce parcours. Je te suis reconnaissant de t'y être engagé directement et d'avoir œuvré à diligenter l'achèvement de cette thèse.

des membres de mon comité de thèse, Dany Laveault, Ph. D. et Liliane Dionne, Ph. D., qui ont élevé le niveau et la qualité de ce travail par leurs précieuses rétroactions. Durant nos rencontres, j'ai vraiment eu le sentiment que chacun de vous voulait me voir produire une thèse qui puisse à la fois ouvrir des pistes intéressantes pour la recherche et être utile pour l'enseignant et le concepteur de guides pédagogiques.

de Carole Fleuret, Ph. D., et Raymond Lebanc, Ph. D., qui par leurs conseils m'ont aidé dans l'organisation de ma thèse et la méthodologie de la recherche. Carole, vos enseignements sur le plan méthodologique m'ont guidé tout au long de cette thèse. Merci.

de mes amis Lilia Forte et Eddy Kamali pour leur participation à la validation de mes instruments de mesure et leurs conseils et encouragements tout au long de ce projet.

de mes amis Jean-Félix Nsegbé, Guy-Maurice Nouwa et Maud Mediel pour les multiples relectures de cet ouvrage. Je vous ai souvent tenus en éveil très tard pour m'aider; en voici le résultat. Merci.

du Bon Dieu qui m'a inspiré durant mon parcours académique. Merci Seigneur, mon Dieu, de ce que don de savoir que tu m'as accordé durant ce parcours. Sans toi rien je ne suis rien. Gloire à toi!

Dédicace

À ma tendre épouse Marlyse,

qui avec une patience déconcertante, a fait sien ce chef d'œuvre et m'a encouragé à poursuivre cette entreprise, surtout quand le vice du désespoir et de l'abandon séduisait et me poussait à rendre les armes.

*À mes enfants : Mike Christian,
 Steven Brian,
 Léana Estelle et
 Cassy Alexandra*

qui ont été source d'inspiration et de confiance.

À ma mère Jeanne Fola,

qui a toujours été présente dans ma vie et m'a soutenu dès le début de ce projet;

À mes frères et sœurs Thérèse Ngoula, Émeline Fola, Rodrigue Fola, Ghislain Fola, Olivier Fola et Tatiana Moukam; mes nièces Nadine Kouchere et Thérèse Ngoufo.

Résumé

Selon plusieurs recherches, le guide pédagogique aide les enseignants à mettre en œuvre au quotidien les prescriptions curriculaires et par ricochet, effectuer des évaluations formatives de haute qualité. Ce niveau de qualité passe par le respect des principes qui sous-tendent l'évaluation formative. Le but de cette recherche était de scruter, à travers une analyse de contenu qualitative, trois guides pédagogiques de sciences et technologie pour voir comment les principes d'évaluation formative seraient pris en compte dans les activités d'évaluation proposées aux enseignants à des fins formatives. Nous avons conçu et validé deux instruments de collecte de données à l'aide d'un système de codage basé sur les dimensions qui ont émergé des principes d'évaluation formative. Les résultats révèlent que les activités proposées ne se conforment pas toujours au processus d'évaluation formative et aux conditions entourant leur mise en œuvre. Plusieurs implications théoriques et pratiques émanant de cette étude vont permettre d'étayer davantage le sujet et également inspirer les enseignants dans leurs démarches d'évaluation formative.

Abstract

Many researchers have recognized the teacher's textbook as an appropriate resource that could help teachers to fulfill curricular requirements in their challenge to perform high quality formative assessment. This could be possible if formative assessment principles were integrated in each formative assessment activity. This study aims to examine, through content analysis, grade 8 science and technology teacher's textbooks to see how formative assessment principles were taken into account in suggested assessment activities intended for formative purpose. We designed two data collection instruments based on the formative assessment principles. Then, we explored the teacher's textbooks with the instruments and analyzed the data. The results tend to indicate that formative assessment principles were not always integrated into the formative assessment process and the conditions surrounding formative assessment activities. Theoretical and practical implications from this study should offer an avenue of research in this field and also inspire teachers in their formative assessment activities.

Table des matières

Remerciements.....	ii
Dédicace	iii
Résumé	iv
Abstract	iv
Liste des tableaux.....	vi
Liste des annexes	vi
Liste des abréviations.....	vi
Chapitre 1 : Introduction générale	1
1.1. Problématisation de l'étude	1
1.2. Conceptualisation de l'étude	8
1.2.1. Perspective épistémologique	8
1.2.2. Dimensions et concepts associés aux principes d'évaluation formative.....	8
1.2.3. Question de recherche	12
1.3. Opérationnalisation de l'étude	12
Chapitre 2 : Article	21
2.1. Introduction	21
2.2. Recension des écrits	24
2.2.1. Rôle du guide pédagogique	25
2.2.2. Principes d'évaluation formative	26
2.3. Cadre conceptuel	28
2.2.1. Processus	28
2.2.2. Conditions	29
2.4. Méthodologie	30
2.2.1. Unités d'étude et unités d'analyse.....	30
2.2.2. Catégories analytiques et système de codage.....	31
2.2.3. Instruments	32
2.2.4. Procédure.....	33
2.5. Résultats, interprétation et discussion	33
2.6. Limites et conclusion	50
Chapitre 3 : Conclusion générale.....	53
3.1. Rappel du contexte.....	53

3.2.	Retombées de l'étude	55
3.2.1.	Implications théoriques	55
3.2.2.	Implications pratiques	56
3.3.	Limites.....	56
	Références.....	58
	Annexes	64

Liste des tableaux

Tableau 1 :	Occasions d'évaluations dans Perspectives	15
Tableau 2 :	Catégories analytiques et codes	16
Tableau 3 :	Instrument de collecte de données relatives au processus	17
Tableau 4 :	Instrument de collecte de données relatives aux conditions	18
Tableau 5 :	Quantification des données par dimension – Processus	38
Tableau 6 :	Quantification des données par dimension – Conditions.....	45
Tableau 7 :	Occurrence des données – exemple du GPI.....	49

Liste des annexes

Annexe 1 :	Tableau synthèse des principes d'évaluation formative.....	64
Annexe 2 :	Mise en relation et synthèse des principes	65
Annexe 3 :	Occurrence des données par guide pédagogique - GPP	67
Annexe 4 :	Occurrence des données par guide pédagogique - GPI.....	68
Annexe 5 :	Occurrence des données par guide pédagogique - GPT.....	69
Annexe 6 :	Exemples d'extraits tirés des guides pédagogiques analysés.....	70

Liste des abréviations

GPP : Guide pédagogique Perspectives sciences et technologie 8^e année

GPI : Guide pédagogique Investigation sciences et technologie 8^e année

GPT : Guide pédagogique Technoscience

Chapitre 1 : Introduction générale

La mise en œuvre de l'évaluation formative a toujours été un défi pour les enseignants. Cette étude présente le guide pédagogique comme un des moyens dont les enseignants disposent pour mettre en œuvre l'évaluation formative, sous réserve que ces guides pédagogiques soient eux-mêmes respectueux des principes d'évaluation formative. Cette thèse est rédigée sous forme d'article et comprend trois parties. Dans la première, nous présentons notre étude. Ensuite, la deuxième est consacrée à l'article proprement dit qui, au moment de la publication de la thèse, est prêt à être soumis à une revue scientifique du domaine de la mesure et de l'évaluation. La dernière partie sera la conclusion générale de l'étude.

1.1. Problématisation de l'étude

Le monde de l'éducation a connu depuis quelques décennies d'énormes changements, voire une réforme des politiques et pratiques évaluatives afin d'accroître le rendement scolaire des élèves (Black et William, 1998a; Bransford, Brown et Cocking, 2000; Buchanan, 2000; Henly, 2003). Comme un des éléments-clés de ces réformes, l'évaluation formative s'est démarquée nettement avec une fulgurante progression depuis le milieu des années 1990 au point de métamorphoser les pratiques pédagogiques dans de nombreux pays (Cizek, 2010; Jalbert et Munn, 2001; Mottier-Lopez, 2007). En Amérique du Nord, notamment au Canada, les programmes et les curriculums dans plusieurs provinces ont été modifiés, et ont subi des réformes systémiques afin de s'arrimer à la mouvance. Selon le Conseil des Ministres de l'Éducation du Canada [CMEC] (2004), les réformes scolaires du début du 21^e siècle portent entre autres sur les politiques curriculaires, le contenu des programmes et les stratégies d'enseignement et apprentissage. En Ontario, on parle d'une réforme centrée sur l'élève qui, amorcée vers la fin des années 1990, en est rendue en matière d'évaluation à sa phase de consolidation avec la première édition en 2010 de la politique sur l'évaluation du Ministère de l'Éducation de l'Ontario (MÉO).

Bien que l'évaluation formative aide à éclairer les stratégies d'amélioration dans le système éducatif tant au niveau des classes, au niveau des établissements qu'au niveau des politiques éducatives, sa mise en œuvre connaît encore aujourd'hui des obstacles. Ceux-ci sont liés entre autres à la sous-représentativité des principes d'évaluation formative dans la conceptualisation de l'évaluation formative, surtout dans le domaine des sciences, qui comparativement aux autres domaines d'études comme la littérature et la numératie, a fait l'objet d'un peu moins d'intérêts sur

le plan de la recherche (Bennett, 2011; Black et Wiliam, 1998b; Kesidou et Roseman, 2002; Lin, Chang et Cheng, 2010; OCDE, 2005; OCDE/CERI, 2008). Cette mise en œuvre est assez exigeante, quel que soit le domaine, surtout pour l'enseignant qui devrait s'y engager activement en ciblant principalement les apprentissages de l'élève, plutôt que d'y adhérer en apparence par la mise en œuvre des « pratiques mécaniques et superficielles » qui n'ont rien à voir avec l'évaluation formative à proprement parler (Allal et Laveault, 2009, p. 101 ; Halverson, Grigg, Prichett, et Thomas, 2007).

Pour ce faire, le MÉO met un accent sur les échanges autour de la pratique éducative, les efforts collectifs, la collaboration et le soutien mutuel du personnel enseignant dans le cadre des communautés d'apprentissage professionnelles (CAP) comme axe du développement professionnel des enseignants. La politique du MÉO sur l'évaluation vient donc en renfort pour guider et appuyer les efforts des enseignants dans leurs diverses pratiques évaluatives. Cette démarche est soutenue par plusieurs recherches qui montrent que la participation des enseignants de sciences aux activités de développement professionnel les aide à intégrer l'évaluation formative dans leur enseignement (Feldman et Capobianco, 2008 ; Torrance et Pryor, 2001). Pour Black et Wiliam (1998b), les enseignants ont besoin de soutien professionnel et d'outils afin de mettre en œuvre des évaluations de haute qualité. Ce niveau de qualité ne peut toujours pas être atteint par les seules CAP en ce sens que les échanges effectués entre les enseignants n'aident que progressivement ceux-ci à forger et consolider leur expérience en évaluation formative. Comme autre support pédagogique dont les enseignants disposent pour appuyer leurs pratiques évaluatives, il y a le guide pédagogique.

Le guide pédagogique est un « document pédagogique conçu comme complément à un programme d'études afin d'en faciliter la compréhension, d'en favoriser l'application et de susciter l'innovation pédagogique » (Legendre, 2005, p. 728). Il a pour but de proposer aux enseignants des démarches pour un suivi efficace du curriculum, une aide pour maintenir le fil du cours sans s'écarter des objectifs tout en respectant au besoin les étapes bien précises du déroulement des activités de classe ou de l'unité de cours. Ainsi, le guide pédagogique devrait appuyer l'action de l'enseignant en proposant des stratégies d'évaluation qui occupent pour ainsi dire une place importante dans la planification de l'apprentissage (MÉO, 2010). Le guide pédagogique est très utilisé par les enseignants de façon générale, et plus particulièrement ceux du domaine des sciences et technologie, pour la préparation et la mise en œuvre de leur enseignement (Bucheton, 1999 ;

Hasni, Moresoli, Samson et Owen, 2009 ; Margolinas et Worzniak, 2009). Le guide pédagogique est un moyen pour offrir à l'enseignant l'information mise à jour, tant sur le contenu et la matière que sur les méthodes d'enseignement et d'évaluation nouvelles ou développées. Il est comme une sorte de vecteur très efficace d'actualisation du curriculum, lorsque la mise à niveau des enseignants traine à suivre l'évolution du curriculum (Nicolot, 2003). Ainsi, pour une mise en œuvre plus intensive de l'évaluation formative, le guide pédagogique devrait aider les enseignants à la mettre en pratique au quotidien.

Rappelons que, de par son rôle, le guide pédagogique devrait aider les enseignants à suivre le curriculum. Ce rôle n'est pas figé parce que selon les politiques éducationnelles, le guide pédagogique peut être prescriptif ou consultatif. Les guides pédagogiques consultatifs sont flexibles et laissent à l'enseignant le libre choix de se servir de ses propres principes pédagogiques pour planifier ses leçons et concevoir ses propres ressources pédagogiques. Quant au guide pédagogique prescriptif, il est structuré et procure à l'enseignant un cadre pédagogique, ainsi que des suggestions et approches pédagogiques afin de l'aider à mieux planifier ses unités de cours (Doyle, 1990). En Ontario, le guide pédagogique serait plutôt consultatif parce que le MÉO met l'accent sur les CAP et laisse à l'enseignant le libre arbitre de choisir parmi les guides pédagogiques approuvés, ceux qu'il juge les plus appropriés pour bâtir ses leçons. Ainsi, l'enseignant ontarien peut se servir de plusieurs guides pédagogiques à la fois, mais toutefois, à titre consultatif. Cependant, les critiques du guide pédagogique consultatif s'appuient généralement sur le fait que les enseignants ne sont pas des spécialistes en tout; en plus des contraintes de temps liées à la conception et la mise en œuvre de matériel didactique. De ce fait, l'enseignant devrait-il concevoir ses propres instruments d'évaluation ou alors se servir des instruments existants et validés par des experts en la matière? Cette question est bien difficile parce qu'on aurait tendance à pencher pour l'usage d'instruments validés. Cependant, bien qu'en contexte socioconstructiviste l'enseignant soit appelé à construire ses propres instruments, il n'en demeure pas moins vrai que la réalité du terrain tranche parfois autrement; de fait, les enseignants ne sont pas toujours des spécialistes en conception d'instruments de mesure, ceci requérant des compétences métrologiques et parfois docimologiques spécifiques pour élaborer et valider lesdits instruments. Beaucoup d'enseignants s'appuient sur le guide pédagogique pour la préparation de leurs leçons et la planification de leurs unités de cours, car ils n'ont pas toujours le temps de le faire à partir de leurs seuls savoirs théoriques, didactiques et pédagogiques. Il leur est par

conséquent difficile de se lancer dans la construction d'instruments d'évaluation; les nouveaux enseignants, parce que peu expérimentés, sont à la merci des ressources pédagogiques, qui devraient être de haute qualité (Black et Wiliam, 1998b). Pour toutes ces raisons, nous sommes d'avis avec Davis (2006) que la conception de ressources pédagogiques requiert une certaine expertise que tous les enseignants n'ont pas forcément. De plus, nous soutenons aussi que les guides pédagogiques, qu'ils soient prescriptifs ou consultatifs, devraient fournir à l'enseignant des suggestions d'évaluation de haute qualité.

Par ailleurs, certains guides pédagogiques regorgent de beaucoup plus de savoirs disciplinaires que de savoir-faire pédagogiques. Selon Lin et ses collègues (2010), les concepteurs des guides pédagogiques du secondaire devraient rendre ces derniers plus pédagogiques, c'est-à-dire, à même de guider les enseignants dans leur pratique enseignante. De même, comme le soulignent Reys, Reys et Chávez (2004), le guide pédagogique devrait procurer à l'enseignant des activités devant lui permettre de mobiliser les élèves dans leurs apprentissages. Ainsi, un guide pédagogique, du fait qu'il propose aux enseignants des activités d'évaluation formative, devrait faire en sorte que ces activités puissent respecter les principes qui sous-tendent l'évaluation formative de façon à ce que les enseignants puissent en tirer profit au maximum.

Selon McManus (2008), « formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes » (p. 3). L'évaluation formative est donc un processus continu et dynamique qui engage l'élève tout au long de sa mise en œuvre et dans lequel certaines activités entreprises tant par l'élève que par l'enseignant fournissent des informations pouvant mener à une rétroaction. Cette définition de l'évaluation formative distingue celle-ci de l'évaluation sommative en ce que cette dernière n'est pas toujours continue et dynamique, et ne donne pas toujours lieu à des rétroactions qui puissent aider l'enseignant à modifier l'enseignement en cours. À cette définition, on peut associer celle du principe. Un principe peut être défini en sciences de l'éducation comme « une proposition générale ou une norme régie par des fondements théoriques et issue soit de la recherche en éducation, soit de l'expérience des enseignants » (Legendre, 2005, p. 1074). En d'autres termes, on pourrait dire que les théories sur lesquelles s'appuie l'évaluation formative sont constituées de principes. Dans les lignes ci-après, nous énumérons les listes de principes trouvés dans les écrits. Ces listes ont été ajustées de façon à ne retenir que les principes exclusivement rattachés à l'évaluation formative.

En 2005, l'OCDE a mené des études de cas à l'échelle internationale dans des pays où l'évaluation formative a fait ses preuves. De l'ensemble de ces études, émanent cinq principes fondamentaux de l'évaluation formative que les enseignants devraient appliquer afin de créer un environnement d'enseignement et d'apprentissage dynamique et faire progresser les élèves vers les objectifs d'apprentissage préétablis tout en développant des aptitudes de « savoir apprendre ». Nous retenons trois de ces principes en l'occurrence ceux en lien avec l'évaluation formative :

1. instaurer une culture de classe qui encourage l'interaction et l'utilisation des outils d'évaluation. L'environnement de la classe doit être propice aux échanges et à l'emploi d'outils d'évaluation. Ceci contribuerait à sécuriser les élèves de façon à favoriser la prise de risques et les amènerait à plus d'audace et de courage dans leurs interventions. Dans un tel climat, les élèves peuvent aisément faire des erreurs constructives;
2. rétroagir sur les performances de l'élève et adapter l'enseignement pour répondre aux besoins identifiés. La rétroaction doit intervenir au moment opportun, être précise et comporter des conseils pour améliorer les performances futures, mais aussi se rattacher à des critères explicites relatifs aux performances attendues;
3. impliquer activement les élèves dans le processus d'apprentissage en général, afin de les aider à développer des stratégies métacognitives et des techniques d'autoévaluation.

Une autre liste de principes est proposée par Clark (2010), dans un article de recherche théorique sur l'évaluation formative en contexte de réforme. Selon cet auteur, sept principes clés énumérés ci-après sont essentiels à l'établissement d'un lien entre l'enseignement, le curriculum et l'évaluation formative, dans l'optique d'appuyer les apprentissages individuels en contexte de salle de classe :

1. engager les élèves à comprendre clairement ce qu'ils apprennent et ce qu'on attend d'eux;
2. donner aux élèves une rétroaction immédiate sur la qualité de leur travail et sur les possibilités d'amélioration;
3. prodiguer aux élèves des conseils pour soutenir leur progrès;
4. impliquer les élèves dans les décisions relatives aux étapes subséquentes;
5. informer les élèves sur la disponibilité du soutien et leur en procurer en cas de besoin;
6. engager les élèves à construire leurs propres connaissances et à réfléchir sur leurs apprentissages;

7. amener les élèves à être plus responsables de leurs apprentissages et participer au processus d'apprentissage.

Le Comité consultatif mixte (1993) a élaboré des principes relatifs aux pratiques d'évaluation des apprentissages scolaires en contexte de salle de classe au Canada, chacun de ces principes étant suivi d'une série de lignes directrices. Bien que ces lignes directrices soient aussi pertinentes les unes que les autres dans l'ensemble, nous avons retenu celles qui portent sur l'évaluation formative :

1. les méthodes d'évaluation devraient être adaptées aux buts de l'évaluation et à son contexte général. Il s'agit de lier les méthodes d'évaluation aux buts et aux objectifs d'apprentissage ainsi qu'au bagage personnel des élèves et à leurs expériences. Tout langage susceptible de choquer un groupe particulier d'élèves devrait être proscrit;
2. on devrait offrir à tous les élèves suffisamment d'occasions de manifester les connaissances, les habiletés, les attitudes et les comportements qui font l'objet de l'évaluation. De ce fait, il faudrait informer l'élève sur l'intention d'évaluation et l'encourager à répondre à tous les items sans craindre d'être pénalisé;
3. les rapports d'évaluation devraient être clairs, précis et avoir une valeur pratique pour les personnes à qui ils s'adressent. Par exemple, une description correcte des forces et des faiblesses de l'élève lors de la communication permet de limiter l'erreur systématique.

Selon Trumpower et Sarwar (2010), le processus d'évaluation formative devrait respecter les quatre principes suivants :

1. permettre de mesurer les apprentissages pertinents. Une évaluation formative valide devrait être plus orientée et focalisée sur les habiletés supérieures, transférables et conceptuelles, plutôt que de porter principalement sur la mémorisation;
2. identifier les forces et les faiblesses de l'élève afin de permettre à l'enseignant et à l'élève lui-même de préparer une rétroaction individualisée pour remédier aux écarts observés et améliorer l'apprentissage;
3. fournir une rétroaction efficace afin de guider l'élève dans ses objectifs d'amélioration de ses apprentissages;

4. être simple d'utilisation. L'efficacité et la survie de l'évaluation formative, au regard de la perception de lourdeur et surplus de travail qu'elle semble engendrer tant chez l'élève que l'enseignant, résident dans sa simplicité.

Une autre série de principes nous est présentée par McMillan (2004) dans un ouvrage portant sur l'évaluation de classe. Ces principes consistent à :

1. établir une évaluation de classe de haute qualité en ce qui concerne la justesse des méthodes d'évaluation, la validité, la fidélité, l'impartialité, etc. La haute qualité d'une évaluation est définie par sa capacité à satisfaire aux spécifications de validité, fidélité et efficacité;
2. évaluer le progrès des élèves pendant les unités de cours;
3. effectuer de l'évaluation de performance à l'aide du dossier d'apprentissage (portfolio) afin de mesurer les habiletés de compréhension profonde et de la pensée critique.

Cizek (2010) quant à lui propose cinq principes devant guider les activités d'évaluation formative. Il s'agit :

1. de faire en sorte que l'élève prenne en charge son apprentissage;
2. d'encourager les élèves à veiller sur leurs propres progressions vers les objectifs d'apprentissage;
3. de fournir des évaluations fréquentes, y compris les évaluations par les pairs et les autoévaluations et des évaluations intégrées aux activités d'apprentissage;
4. d'inclure de la rétroaction qui soit spécifique, opportune, axée sur les objectifs d'apprentissage et non évaluative, et qui puisse fournir à l'élève des occasions de réviser et améliorer son travail, puis approfondir sa compréhension;
5. de promouvoir la métacognition et la réflexion de l'élève sur son propre travail.

Aux travaux des précédents auteurs qui ont dressé des listes de principes, nous pouvons ajouter Scallon (1988a) qui est un auteur clé du domaine de l'évaluation formative. Selon lui, les scénarios d'évaluation formative devraient être souples afin d'assurer « la régulation de l'apprentissage par une rétroaction qualitative » (p. 163), contrairement à la rigidité du scénario d'évaluation sommative qui permet de garantir « la crédibilité de l'évaluation et augmenter sa fiabilité (objectivité) » (p. 163).

Les principes d'évaluation formative qui viennent d'être énumérés pourraient permettre d'améliorer les apprentissages des élèves, si ces principes sont pris en compte dans des activités d'évaluation prescrites dans les guides pédagogiques à des fins formatives. Nous pensons avec Lynch, Kuipers, Pyke et Szesze (2005) et Niclot (2003) que le guide pédagogique est un moyen dont l'enseignant peut se servir pour mettre en œuvre l'évaluation formative. Cette étude a comme objectif d'examiner l'articulation des principes d'évaluation formative dans les guides pédagogiques de sciences. En d'autres termes, nous tentons de répondre à la question générale de recherche suivante : *comment le guide pédagogique de sciences et technologie intègre-t-il les principes d'évaluation formative?* Cette question générale de recherche peut sembler large. Aussi nous nous servons du cadre conceptuel ci-après pour l'affiner.

1.2. Conceptualisation de l'étude

Le cadre conceptuel de la présente étude permet de décrire et de situer les concepts clés qui émergent de la recension des écrits en un tout cohérent. Dans cette section, les concepts clés sont analysés et mis en relation les uns avec les autres. Puisqu'il s'agit de la présentation des concepts clés de l'étude, nous pensons qu'il serait opportun de préciser également à cette étape notre posture épistémologique, c'est-à-dire l'angle sous lequel nous percevons notre étude, ce qui est éventuellement déterminant pour l'interprétation des données.

1.2.1. Perspective épistémologique

Cette étude part du postulat que la prise en compte des principes d'évaluation formative dans les guides pédagogiques pourrait aider l'enseignant à la mettre en œuvre efficacement au quotidien. L'analyse de contenu qualitative convient à ce type d'étude parce qu'elle permet d'examiner de quelle manière les principes d'évaluation formative sont pris en compte dans les guides. Dans le cas de notre étude, les données sont de nature qualitative, et l'épistémologie sous-jacente se veut interprétative en ce sens que nous essayons d'apporter un éclairage sur les pratiques éducatives et des façons de faire en matière d'évaluation formative (Savoie-Zajc, 2011). Le regard interprétatif nous aide donc à comprendre le sens des données extraites d'un guide pédagogique et comment celles-ci nous permettent de nous prononcer sur l'intégration des principes d'évaluation formative dans les guides pédagogiques.

1.2.2. Dimensions et concepts associés aux principes d'évaluation formative

Les principes d'évaluation formative émanant de la recension des écrits sont nombreux. Leur regroupement a permis de constater l'émergence de deux importantes dimensions : le processus et les conditions. Premièrement, comme le souligne Popham (2008), l'évaluation formative est un processus et non un test, contrairement à la pensée populaire. Ce processus est constitué d'étapes allant de la formulation des intentions de l'évaluation jusqu'à la fourniture de la rétroaction descriptive à l'élève. Deuxièmement, les conditions de mise en œuvre de l'ensemble de l'activité d'évaluation formative sont régies par des principes. Mentionnons que si certains principes portent sur une seule dimension, plusieurs autres sont multidimensionnels. Par exemple, l'OCDE (2005) stipule qu'il faudrait instaurer une culture de classe qui encourage l'interaction et l'utilisation des outils d'évaluation. Ce principe porte sur l'« interactivité » (condition) en ce sens qu'il suggère une interaction entre les élèves. Il porte également sur la « mesure » dans le processus parce qu'il incite à l'utilisation des outils de mesure. Par ailleurs, les principes associés à la même dimension sont identifiés par le terme « concept ». Ainsi, la mesure est un concept relié au processus, de même que l'interactivité est un concept relié aux conditions.

Le processus d'évaluation formative, comme toute démarche évaluative, est constitué d'étapes. La présence de chaque étape a une importance capitale pour l'ensemble du processus. Par exemple, une activité d'évaluation formative qui ne comporte ni régulation ni mécanisme d'auto-observation de la part de l'élève n'apporte pas grand-chose aux apprentissages de l'élève. Forgette-Giroux et Simon (1998) proposent quatre étapes :

1. intention : l'intention est une étape de planification qui consiste à établir :
 - a) la fonction de l'évaluation, c'est-à-dire le but ou le pourquoi de l'évaluation,
 - b) le responsable d'évaluation, c'est-à-dire la personne qui évalue,
 - c) l'objet de l'évaluation, ou le quoi, en d'autres termes il s'agit de mentionner ce sur quoi va porter l'évaluation;

Notons que l'étape d'intention aura trois concepts : l'intention - fonction de l'évaluation; l'intention - responsable d'évaluation et l'intention - objet de l'évaluation;

2. mesure : lors de cette étape, il est question de déterminer la tâche et l'instrument de cueillette d'informations appropriées qui permettront de mesurer, observer ou capter le niveau de rendement démontré par l'élève. L'étape de mesure aura deux concepts : la mesure – tâche et la mesure – instrument;

3. jugement : le jugement permet de situer le rendement que l'élève a démontré à l'étape de mesure, selon une règle, un critère ou une norme établie. Les critères de réussite ou critères d'évaluation sont spécifiés par souci d'interprétation commune. Ces critères résultent de l'opérationnalisation de l'objet d'évaluation en indices de performance établis lors de l'élaboration de l'intention, et à partir desquels s'appuie le jugement. À cette étape, il convient d'avoir des outils de jugement comme des échelles d'appréciation sur lesquels sont souvent consignés les critères de réussite ainsi que les règles d'interprétation. L'étape de jugement aura deux concepts : le jugement – outils et le jugement - critères d'évaluation;
4. décision et communication : cette étape consiste à choisir entre différentes actions permettant de réajuster la démarche d'apprentissage de l'élève à travers la régulation de l'enseignement et de l'apprentissage (Allal, 1993; Forgette-Giroux et Simon, 1998). La communication dans une activité d'évaluation formative est toujours orientée vers la rétroaction afin d'apporter des modifications éventuelles au processus d'enseignement et d'apprentissage dans lequel l'enseignant et l'élève sont engagés (Black, Harrison, Lee, Marshall et Wiliam, 2004; Scallon, 2000). Les nombreux travaux autour de la rétroaction font ressortir certaines caractéristiques de la rétroaction : clarté, descriptivité, précision, opportunité, ponctualité, fréquence, signifiante, etc. qu'il convient de retrouver dans une rétroaction en contexte formatif (Allal et Mottier Lopez, 2005 ; Brookhart, 2010 ; Cizek, 2010 ; Earl et Katz, 2006 ; Hattie et Timperley, 2007 ; OCDE, 2005 ; Scallon, 1988b).

En outre, comme mentionné plus haut, certains principes d'évaluation formative relèvent des conditions qui entourent les activités d'évaluation formative. Contrairement au processus qui est constitué d'étapes, les conditions consistent en une liste de concepts qu'on pourrait déceler dans une activité d'évaluation formative. Il s'agit des concepts ci-après :

1. souplesse et dynamisme : la souplesse veut que l'enseignant et l'élève aient beaucoup de latitude dans l'application du processus d'évaluation. Par exemple, ils peuvent effectuer des changements aux modalités d'évaluation en cours. Ces scénarios devraient également être dynamiques dans le sens des échanges fructueux élève – enseignant; par exemple en discutant pour s'assurer de la compréhension des critères par les élèves au début de l'évaluation;
2. droit à l'erreur : contrairement à ce que beaucoup d'élèves pensent, avoir droit à l'erreur ne veut pas dire qu'il faut délibérément laisser des erreurs dans son travail pour bénéficier

de la rétroaction. Ce qui est important de savoir c'est que la rétroaction se fait généralement progressivement à l'instar de la zone de développement proximale de Vygotsky (1978) où le niveau de difficulté dans la tâche proposée augmente en fonction des performances démontrées. À cet effet, lorsque l'élève démontre une meilleure performance, il reçoit en retour une meilleure rétroaction. Cependant, les erreurs ne devraient aucunement être utilisées pour sanctionner la compétence de l'élève, mais comme une occasion d'*autorétroaction* significative pour se réajuster par la suite;

3. différenciation : c'est l'adaptation des méthodes d'évaluation au bagage de chaque élève et la promotion de l'égalité de chances d'apprentissage pour tous les élèves (Laveault, 2010);
4. responsabilisation : la responsabilisation permet à l'élève d'autoréguler ses apprentissages en participant à toutes les étapes du processus d'évaluation formative (Laurier, Tousignant et Morissette, 2005; Wiggins, 1993). Ainsi les activités d'autoévaluation responsabilisent l'élève. La responsabilisation fait également appel à la métacognition, car le processus d'évaluation formative vise à développer et à favoriser la métacognition chez les élèves. Selon Earl et Katz (2006), « il y a métacognition lorsque les élèves veillent eux-mêmes à leur apprentissage et qu'ils se servent des rétroactions ainsi recueillies pour faire des ajustements, des adaptations et même des changements importants à ce qu'ils comprennent » (p. 13). La métacognition suscite en outre chez l'élève une évolution du stade d'apprenant autonome à celui d'évaluateur autonome, capable de faire des liens entre l'évaluation et l'apprentissage (Earl et Katz);
5. interactivité : le sens que nous donnons à l'interactivité dans ce contexte serait les échanges entre les élèves en ce qui concerne la participation mutuelle à l'évaluation formative, soit l'évaluation par les pairs. L'enjeu de cette évaluation est double parce que non seulement l'élève évalué se sent proche de son évaluateur, ce qui favorise la prise de risque comme mentionné plus haut, mais en plus, certains problèmes peuvent implicitement être résolus; il arrive par exemple que les évalués soient confrontés aux problèmes d'incompréhension sur la formulation d'une question. Alors, un échange ou une interaction avec les pairs pourrait permettre de lever l'équivoque (Scallon, 2004).

Notons qu'il serait plus commode de parler d'autoévaluation et d'évaluation par les pairs dans le processus étant donné que la responsabilité de l'évaluation en découle naturellement, mais dans le cadre de cette étude, nous préférons les traiter dans les conditions parce que nous les avons respectivement associées à la responsabilisation et l'interactivité.

Nous venons de définir un cadre constitué de concepts clés regroupés selon deux dimensions à savoir le processus d'évaluation et les conditions, développées à la lumière des écrits scientifiques. Ces dimensions encadrent les principes d'évaluation formative que nous avons énumérés dans la recension des écrits aux fins de l'examen des guides pédagogiques. Par ailleurs, les arguments évoqués depuis le début permettent de constater le besoin, sur le plan de la recherche, de mettre en lumière la contribution du guide pédagogique au succès de l'apprentissage de l'élève à travers les activités qu'il propose aux enseignants. Ceci nous permettra de développer notre question de recherche.

1.2.3. Question de recherche

En matière d'évaluation, les réformes scolaires qu'a connues le monde de l'éducation durant ces dernières décennies mettent un accent particulier sur l'évaluation formative des apprentissages (Morissette, 2009). L'évaluation formative s'effectue de plusieurs manières, par exemple par les activités planifiées ou spontanées que l'enseignant effectue avec les élèves et qui sont habituellement proposées dans les guides pédagogiques. Ceux-ci ont plusieurs rôles dont celui de guider les enseignants dans l'intégration des changements curriculaires issus des réformes. Cette guidance se démarque notamment en matière de méthodes d'enseignement et d'évaluation nouvelles, ou en cours d'élaboration, ou encore en fonction de celles qui tardent à s'opérationnaliser dans la pratique enseignante quotidienne, comme ce serait le cas de l'évaluation formative (Beyer, Delgado, Davis et Krajcik, 2009; Lynch *et al.*, 2005; Niclot, 2003).

Par conséquent, il serait d'un certain intérêt de voir comment les principes d'évaluation formative éventuellement présents dans les guides pédagogiques participent à rendre formatives les activités d'évaluation qu'on y retrouve, et ce dans le but d'améliorer l'enseignement. Ce qui nous amène à formuler la question spécifique de recherche comme suit : *De quelle manière les activités d'évaluation proposées dans les guides pédagogiques de sciences et technologie de l'Ontario à des fins formatives prennent-elles en compte des principes d'évaluation formative?* Pour tenter de répondre à cette question spécifique de recherche, nous avons eu besoin d'un cadre méthodologique pour opérationnaliser notre étude.

1.3. Opérationnalisation de l'étude

Nous avons procédé à une analyse de contenu de guides pédagogiques. C'est une opération qui permet d'effectuer une analyse rigoureuse des textes et des messages contenus dans les

documents, en utilisant des techniques scientifiques, objectives et systématiques d'analyse de textes (Richard, 2006; Gohier, 2011).

Les sciences dans le système éducatif de l'Ontario, englobent les sciences et la technologie. Notre intérêt est porté sur la technologie, définie par l'International Technology Education Association [ITEA] (2000), comme l'art de fabriquer, confectionner ou réaliser des choses pour accroître les possibilités humaines et satisfaire en même temps aux besoins et bien être des êtres humains - traduction libre - (p. 2). À cette définition, on peut associer celle du MÉO qui met un accent sur le travail pratique et la réalisation des projets technologiques en vue de produire des biens ou offrir des services (MÉO, 2009). Le MÉO tient à jour une liste de guides pédagogiques approuvés appelée la *Liste Trillium*. Aucun guide pédagogique de technologie de la 9^e à la 12^e année ne figurait dans cette liste au moment de cette étude. Les guides pédagogiques de sciences et technologie de 6^e, 7^e et 8^e année y étaient présents. Le palier intermédiaire, c'est-à-dire les classes de 7^e et 8^e année, constitue un pont entre les principes scientifiques de base du cycle élémentaire et les connaissances scientifiques et technologiques avancées du cycle supérieur. Selon l'American College Testing (2008), il existe un lien entre les performances des élèves en 8^e année et leur choix de carrière postsecondaire. À la fin de la 8^e année, les élèves ontariens sont appelés à choisir des cours collégiaux ou universitaires et ceux qui auront été séduits par la technologie pourront donc suivre des cours de technologie, cours qui pourrait avoir un impact sur leur carrière. Ces caractéristiques nous ont amené à considérer l'analyse des guides pédagogiques de sciences et technologie de la 8^e année.

Au moment de l'élaboration de cette étude, quatre guides pédagogiques de sciences et technologie de 8^e année étaient approuvés dans la Liste Trillium : deux provenant du secteur francophone et deux autres du secteur anglophone. Ces guides pédagogiques proviennent de deux maisons d'édition différentes : Duval et ERPI. À la suite d'une analyse comparative des documents francophones et anglophones que nous avons faite avant la collecte des données, nous nous sommes rendu compte que les ouvrages provenant d'un éditeur donné sont traduits d'une langue à l'autre. Par ailleurs, nous avons aussi constaté une certaine similitude sur les plans organisationnel et structurel entre les guides pédagogiques de chaque éditeur. Par conséquent, nous avons retenu les éditions francophones des guides pédagogiques afin de ne pas collecter les données identiques. Il s'agit des guides pédagogiques Perspectives Sciences et technologie 8 de

Duval (2010) et Investigation 8 Sciences et technologie d'ERPI (2009) que nous abrègerons par la suite GPP et GPI respectivement.

En plus de ces guides pédagogiques approuvés le guide pédagogique Technoscience ci-après abrégé GPT publié en 2001 demeure encore, selon le Centre Franco-Ontarien des Ressources Pédagogiques (CFORP), parmi les plus populaires chez les enseignants (CFORP, communication personnelle, 21 mars 2013). Pour cette raison, il a été ajouté au corpus. En effet, la sélection d'un guide pédagogique antérieur à la publication de la politique ministérielle (2010) trouve son sens dans le fait que les politiques d'évaluation en Ontario ont évolué avec le temps avec les parutions de nombreuses politiques jusqu'à celle de 2010. Les principes d'évaluation formative devraient quand même être perceptibles dans un guide pédagogique paru en 2001. Par ailleurs, les principes ne datent pas d'aujourd'hui parce que Scallon en 1988 en a énuméré certains. De plus, la popularité d'un tel ouvrage en fait un sujet d'étude pertinent.

Bien que les deux premiers guides pédagogiques aient été édités en 2009 et 2010, soit avant et pendant l'année de publication de la politique d'évaluation de l'Ontario, nous pensons qu'en raison de leur statut de ressources approuvées par le MÉO, ils devraient refléter les avancées des connaissances en matière d'évaluation formative. De plus, comme nous l'avons souligné dans l'introduction, la réforme en Ontario est passée par plusieurs étapes jusqu'à sa consolidation; son reflet pourrait être perceptible dans des guides édités en 2010 ou 2009. Nous nous retrouvons donc avec un corpus de trois guides pédagogiques : le GPP, le GPI et le GPT. Le curriculum de sciences et technologie de l'Ontario de 2007 est subdivisé en quatre domaines : Systèmes vivants; Matière et énergie; Structures et mécanismes et Systèmes de la terre et de l'espace. Ces domaines comportent chacun un sujet à l'étude. Parmi les sujets à l'étude pour la 8^e année, seul Les systèmes en action (correspondant au domaine Structure et mécanismes) rejoint les concepts technologiques du curriculum de l'Ontario 9^e et 10^e année en éducation technologique. Étant donné que nous sommes intéressé à la technologie, les unités d'études porteront sur le domaine Structure et mécanismes des trois guides pédagogiques retenus.

Il y a aussi lieu de distinguer dans chaque guide pédagogique l'existence de deux sources de données qui sont le préambule et le contenu : premièrement, il y a des données génériques affectant l'ensemble du guide qui se trouvent dans le préambule. Par exemple, on peut lire dans le préambule du GPI : « les enseignants doivent indiquer avec précision aux élèves les éléments qu'ils doivent améliorer et proposer des moyens pour faire progresser ceux d'entre eux qui ont de

la difficulté » (GPI, 2009, p. VE-43). Cet extrait pourrait s'appliquer à l'ensemble du guide. Deuxièmement, il y a des données individuelles à chaque activité qui se retrouvent dans le contenu du guide.

Les unités d'analyse sont constituées d'entités formellement identifiées dans les guides pédagogiques comme étant des activités d'évaluation formative. À titre d'exemple, on peut lire dans le préambule du GPP :

« L'évaluation formative permet au personnel enseignant de vérifier au fur et à mesure le rendement des élèves par rapport aux attentes du curriculum. Le personnel enseignant peut utiliser les résultats de l'évaluation formative pour différencier ses pratiques d'enseignement de façon à mieux refléter les besoins des élèves en matière d'apprentissage. (...) Plusieurs occasions d'évaluation formative sont intégrées dans *Perspectives* sous forme de tâches de performance (...), d'évaluation de la communication orale (...) et de travaux écrits (...) *Perspectives* fournit aux élèves bon nombre d'occasions d'autoévaluation, d'évaluation par leurs pairs et de réflexion sur leur apprentissage (...) Le tableau 1 [*même titre dans le document original*] présente les diverses occasions d'évaluation fournies dans *Perspectives*.» (GPP, 2010, p. xiv).

Dans le « tableau 1 » auquel renvoie l'extrait précédent, on identifie formellement les occasions d'évaluation formative, c'est-à-dire en d'autres mots les activités d'évaluation formative. Ce tableau est inséré dans ce texte à titre d'illustration (Tableau 1). Ceci étant, toutes les activités proposées dans la deuxième colonne du Tableau 1 sont considérées comme des activités d'évaluation formative. Étant donné que dans cette colonne se trouvent des activités, des rubriques et parfois des sections entières, la notion d'activité d'évaluation formative est élargie à toutes les occasions d'évaluation formative du Tableau 1. Dans l'ensemble, on dénombre 40 activités d'évaluation formative dans le GPP, 21 dans le GPI, et 2 dans le GPT.

Tableau 1 : Occasions d'évaluations dans *Perspectives*

Occasions d'évaluation diagnostique Évaluation pour l'apprentissage	Occasions d'évaluation formative Évaluation pour et en tant qu'apprentissage	Occasion d'évaluation sommative Évaluation de l'apprentissage
<ul style="list-style-type: none"> • Aperçu de l'unité • Découvre les sciences et la technologie • Point de départ • Question clé, À voir, Vocabulaire (dans l'Amorce du chapitre) • Histoire de sciences et de technologie 	<ul style="list-style-type: none"> • Sciences en action • Prononce-toi sur un enjeu • Résous un problème technologique • Mène une expérience • Sciences appliquées, Géniales, les sciences! Info techno 	<ul style="list-style-type: none"> • Activité de fin d'unité

<ul style="list-style-type: none"> • Sciences en action • Prononce-toi sur un enjeu • Résous un problème technologique • Réalise une activité • Mène une expérience • Exemple de problème • Sciences appliquées, Géniales, les sciences! Info techno • toutes les photos et illustrations ainsi que les autres éléments visuels 	<ul style="list-style-type: none"> • toutes les photos et illustrations ainsi que les autres éléments visuels • Vérifie ta compréhension • En résumé • Révision du chapitre • Activité de fin d'unité • Fais un résumé • Révision de l'unité
---	---

En ce qui concerne les catégories analytiques, une étude menée en 2012 par Agiro nous a inspiré. Le but de ladite étude était de comprendre comment le guide pédagogique de littérature américaine des écoles secondaires chrétiennes et le manuel de l'élève qui l'accompagne prenaient en compte les aspects tels que la race, l'ethnie, le genre, la classe sociale et les aptitudes physiques et mentales. Chacun de ces aspects ou combinaison de plusieurs aspects a fait l'objet des catégories analytiques. De même, dans la présente étude, chaque dimension a été considérée comme catégorie analytique et chaque concept clé comme code. Le Tableau 2 fait le lien entre les catégories analytiques et les codes.

Tableau 2 : Catégories analytiques et codes

Catégorie analytique	Code
Processus	<ol style="list-style-type: none"> 1. Intention – fonction d'évaluation 2. Intention – responsable d'évaluation 3. Intention – objet d'évaluation 4. Mesure – tâche 5. Mesure – instrument 6. Jugement – outils 7. Jugement – critères d'évaluation 8. Communication – rétroaction
Conditions	<ol style="list-style-type: none"> 1. Souplesse 2. Dynamisme 3. Droit à l'erreur 4. Différentiation 5. Interactivité

Bien que les codes aient été prédéterminés avant la collecte de données, nous sommes resté ouvert à l'émergence d'autres catégories et codes durant la collecte de données. Les catégories analytiques découlent donc des deux dimensions présentées antérieurement. Étant donné la structure de chaque dimension – par exemple les étapes consécutives du processus – et la disparité des deux dimensions – par exemple le processus est de nature séquentielle et les qualités de nature combinatoire –, les instruments de collecte de données ont été conçus de façon à correspondre à chacune des dimensions. Les deux instruments de collectes de données sont les suivants : l'instrument de collecte de données relatives au processus présenté au Tableau 3 et l'instrument de collecte de données relatives aux conditions présenté au Tableau 4. Ces instruments reprennent la synthèse des principes de l'Annexe 2, lesquels principes sont transformés en questions guides de collecte de données auxquelles nous avons essayé de trouver des réponses, en puisant dans les guides pédagogiques des extraits aux fins d'analyse. Ces instruments ont été éprouvés par des exercices de validité qui sont explicités un peu plus loin.

La grille d'analyse des données provenant de chaque instrument comprend en en-tête les unités d'analyse (activité d'évaluation formative) et dans les lignes, les données recueillies dans la colonne « Données provenant du document analysé avec justificatifs ». L'analyse des données porte d'une part sur l'occurrence des principes, par exemple on cherche à savoir si le principe est présent ou absent. Et d'autre part, sur la qualité de la prise en compte du principe; pour ce faire, on se sert des attributs suivants : complète, suffisante, cohérente, pertinente, à jour, désuète, explicite, représentative, etc.

Tableau 3 : Instrument de collecte de données relatives au processus

Document analysé :	Date :	
Titre de l'activité :	Page :	
Activités :	Responsable :	
Code	Question guide de collecte de données	Données provenant du document analysé avec justificatifs
1- Intention	A1. Quels extraits précisent la fonction d'évaluation? Ex. : « cette évaluation est formative » (GPx, 20xx, p. xx).	
	A2. Quels extraits précisent le responsable de l'évaluation? Ex. : « j'évaluerai tes habiletés à communiquer oralement tes idées » (GPx, 20xx, p. xx).	
	A3. Quels extraits précisent l'objet d'évaluation? Ex. : « tu me montreras comment tu mets en application le processus de design » (GPx, 20xx, p. xx).	

2- Mesure	A4. Quelles sont les tâches utilisées ou proposées? Ex. : « construire un objet technologique qui réponde aux exigences du client tout en respectant les normes de sécurité » (GPP, 2010, p. 24).
	A5. Quels sont les instruments utilisés ou proposés? Ex. : « la grille d'observation de la page suivante sera utilisée pour évaluer la mise en application de ton système de levage » (GPx, 20xx, p. xx).
3- Jugement	A6. Quels sont les outils de jugement proposés? Ex. : « Vous pouvez utiliser la grille d'évaluation A pour évaluer les habiletés des élèves dans cette activité » (GPx, 20xx, p. xx).
	A7. Quels sont les critères d'évaluation (critères de réussite) spécifiés? Ex. : « mise en application concrète des résultats d'une activité » (GPx, 20xx, p. xx).
4- Communication	A8. Que propose le guide par rapport à la rétroaction? Ex. : « C'est une occasion de fournir une rétroaction descriptive aux élèves » (GPx, 20xx, p. xx).

Tableau 4 : Instrument de collecte de données relatives aux conditions

Document analysé :	Date :	
Titre de l'activité :	Page :	
Activités :	Responsable :	
Code	Question guide de collecte de données	Données provenant du document analysé avec justificatifs
1- Souplesse	B1. Quels extraits réfèrent à la marge de manœuvre (liberté d'action) accordée à l'élève et l'enseignant par rapport à la mise en œuvre de l'évaluation? (Changements autorisés – Flexibilité quant au moment, lieu, moyens, souplesse dans les changements en cours de route, etc.). Ex. : « modifier les critères d'évaluation en cours de route » (GPx, 20xx, p. xx).	

2- Dynamisme	B2. Quels extraits réfèrent aux interactions entre l'enseignant et les élèves ? Ex. : « si les élèves font l'activité individuellement, discuter des réponses avec la classe » GPx, 20xx, p. 59)
3- Droit à l'erreur	B3. Quels extraits offrent la possibilité à l'élève de dédramatiser les erreurs et en tirer profit? Ex. « S'il y a des équipes qui font des erreurs invitez les élèves à en discuter entre eux avant de poursuivre » (GPx, 20xx, p. xx).
4- Différentiation	B4. Quelles sont les adaptations proposées pour répondre aux besoins particuliers de certains groupes d'élèves ? (prise en compte des élèves avec besoins spéciaux lors de l'évaluation). Ex. : « Les élèves en français langue seconde devront utiliser le tableau 1 afin d'avoir des chances de réussite égales » (GPx, 20xx, p. xx).
5- Responsabilisation	B5. Quels extraits réfèrent à la participation active des élèves aux différentes étapes d'évaluation? [- Implication dans les décisions concernant leur progression; - implication des élèves à veiller sur leurs apprentissages (métacognition)]. Ex. : « Distribuer l'Outil d'autoévaluation et demander aux élèves de s'autoévaluer » (GPP, 20xx, p. XX).
6- Interactivité	B6. Quels extraits favorisent les échanges entre les élèves durant les activités d'évaluation (évaluation par les pairs)? Ex. : « Demandez aux élèves d'échanger leurs copies et s'évaluer mutuellement » (GPx, 20xx, p. XX).

Le premier test de validité a été de conformer les instruments aux règles d'élaboration des listes de vérification. Tout d'abord, ces instruments comportent au plus dix points de vérification (codes) comme requis dans les règles. De plus, chaque principe véhicule une seule idée. Ensuite, les questions redondantes ont été éliminées et celles retenues ont été reclassées de façon à correspondre aux principes préalablement synthétisés. Ceci a permis de rendre les codes intelligibles et mutuellement exclusifs.

Ensuite, un exercice d'appariement des instruments a été effectué; ayant pris soin de masquer les codes, nous avons alors demandé à une collègue en mesure et évaluation, et sans lui montrer l'Annexe 2, d'associer les questions guides des instruments de collecte de données aux codes. Tous les codes ont été appariés correctement.

Ensuite, une étude pilote a été menée afin de vérifier l'exclusivité mutuelle des codes et d'établir un indice de fidélité interjuge. Lors de cette étude pilote, un chercheur autre que l'auteur a été ajouté, dans le but d'augmenter la fiabilité et la crédibilité de l'étude (L'Écuyer, 1987). Ce deuxième chercheur est un collègue qui a reçu une formation de manière à pouvoir mener une

collecte de données de façon autonome. Puis, nous avons choisi au hasard un des guides pédagogiques retenus et nous avons sélectionné au hasard quelques sections que nous avons ensuite balayées, chacun de son côté à l'aide des instruments de collecte de données. Nous avons alors colligé les résultats de cette collecte de données préliminaire et les avons comparés. Étant donné que nous sommes en phase de validation d'instrument et que chaque code renseigne sur la présence ou l'absence des principes d'évaluation formative, les accords sur les occurrences et sur les non-occurrences des principes sont nécessaires. À cet effet, Hawkins et Dotson (1975, cité par House, House et Campbell, 1981) proposent d'utiliser la formule suivante pour calculer la moyenne (M) du pourcentage d'accords sur les occurrences et sur les nonoccurrences : $M(\%) = (A / (A + B + C) + D / (B + C + D)) / 2 \times 100$, où A et D sont les nombres d'accords sur les occurrences et nonoccurrences respectivement, puis B le nombre de désaccords où le chercheur 1 a trouvé des principes et le chercheur 2 n'a rien trouvé, et C le nombre de désaccords où le chercheur 2 a trouvé des principes et le chercheur 1 n'a rien trouvé. Nous avons obtenu une moyenne de 77 %, ce qui selon House, House et Campbell est acceptable. Notons que la modification du nombre d'instruments n'a pas affecté les exercices de validité.

À la suite de l'étude pilote, nous avons procédé à la collecte de données. Les préambules des guides pédagogiques ont été examinés au regard de principes d'évaluation formative synthétisés dans le cadre conceptuel. Lors de cet examen, nous avons regardé la structure de chaque document, comment est conçue l'évaluation formative et les activités d'évaluation identifiées comme formatives que nous avons indiquées dans les unités d'étude. Ensuite, à l'aide de chacun des instruments de collecte de données, nous avons balayé les unités d'analyse en prenant soin de noter dans les instruments de collecte de données, tout extrait pertinent répondant aux questions guides de la collecte de données.

Chapitre 2 : Article

Analyse des principes d'évaluation formative dans des extraits de trois guides pédagogiques de sciences et technologie de 8^e année en Ontario

2.1. Introduction

L'évaluation formative occupe une place prépondérante dans la mise en œuvre des réformes scolaires à l'échelle internationale (Herman, Osmundson et Silver, 2010 ; OCDE 2005). Depuis les débuts des réformes en éducation, tous les domaines éducatifs ont été enrichis des fruits des avancées de la recherche en évaluation des apprentissages, mis à part celui des sciences, qui comparativement aux autres, a fait l'objet d'un peu moins d'attention (Kesidou et Roseman, 2002 ; Lin, Chang et Cheng, 2010). En outre, malgré les proportions que l'évaluation formative prend et la notoriété qui lui est attribuable, il est aberrant d'entendre encore aujourd'hui certains professionnels de l'éducation en sciences associer l'évaluation formative à « l'évaluation qui ne compte pas ». Dérapage ou manquement?

Selon le Conseil des Ministres de l'Éducation du Canada (2004), les politiques curriculaires, le contenu des programmes et les stratégies d'enseignement et apprentissage sont au cœur des réformes scolaires du début du 21^e siècle. En Ontario, on parle d'une réforme centrée sur l'élève, qui a été amorcée vers la fin des années 1990 et qui en est rendue à sa phase de consolidation en matière d'évaluation. Sur le plan du développement professionnel, le ministère de l'Éducation de l'Ontario (MÉO) met l'accent sur les échanges autour de la pratique éducative, les efforts collectifs, la collaboration et le soutien mutuel du personnel enseignant dans le cadre des communautés d'apprentissage professionnelles (CAP). Plusieurs recherches montrent que la participation des enseignants de sciences aux activités de développement professionnel les aide à intégrer l'évaluation formative dans leur enseignement (Feldman et Capobianco 2008 ; Torrance et Pryor 2001). Par ailleurs, Black et Wiliam (1998b) soutiennent que les enseignants ont besoin de soutien professionnel et d'outils afin de mettre en œuvre des évaluations de haute qualité. Au rang de ces outils figure le guide pédagogique. Le guide pédagogique est défini comme un « document pédagogique conçu comme complément à un programme d'études afin d'en faciliter la compréhension, d'en favoriser l'application et de susciter l'innovation pédagogique » (Legendre, 2005, p. 728). Le guide pédagogique reste un outil à la portée de l'enseignant au quotidien, susceptible d'aider ce dernier lors de ses interventions auprès des élèves. De plus, il est

toujours très utilisé par les enseignants (Bucheton, 1999 ; Hasni, Moresoli, Samson et Owen, 2009 ; Margolinas et Worzniak, 2009). C'est sans doute pour cela que le MÉO offre aux enseignants le choix d'utiliser plusieurs guides pédagogiques approuvés pour organiser leurs unités de cours. En ce qui concerne les sciences, le manque d'attention pour cette matière est aussi perceptible en Ontario. Dans les domaines du français et des mathématiques, MÉO a entrepris depuis le début de la réforme en Ontario, plusieurs initiatives comme la création en 2004 du Secrétariat de la littératie et de la numératie dans le but de stimuler le rendement des élèves. Cependant, le domaine des sciences et technologie n'a pas été bonifié d'autant d'intérêts, la technologie encore moins, car lors de notre recherche documentaire, en utilisant des moteurs de recherche tels qu'ERIC, Erudit, etc., nous n'avons pas trouvé d'études portant sur l'évaluation formative et l'éducation technologique en Ontario. Précisons que la technologie dans le cadre de cette étude est le contenu d'enseignement, c'est-à-dire la matière enseignée, basée sur le travail pratique et la réalisation des projets technologiques en vue de produire des biens ou offrir des services (MÉO, 2009). Ceci est à distinguer de la technologie dans le sens de l'usage des Technologies de l'Information et de la Communication (TIC) pour enseigner.

Les guides pédagogiques en usage dans les écoles ontariennes se trouvent dans un document intitulé la *Liste Trillium*, que le MÉO tient à jour. Au moment de cette étude, aucun guide pédagogique de technologie ne figurait dans la *Liste Trillium*. Les guides pédagogiques de sciences et technologie de 6^e, 7^e et 8^e année y étaient présents. De fait, les sciences dans les classes de la 1^{re} à la 8^e année incluent la technologie. Par ailleurs, le palier intermédiaire, c'est-à-dire les classes de 7^e et 8^e année, est un pont entre les principes scientifiques de base du cycle élémentaire et les connaissances scientifiques et technologiques avancées du cycle supérieur. De plus, selon l'American College Testing (2008), il existe un lien entre les performances des élèves en 8^e année et leur choix de carrière au collégial. À la fin de la 8^e année, les élèves ontariens sont appelés à choisir des cours collégiaux ou universitaires et ceux qui auront été séduits par la technologie pourront donc suivre des cours de technologie, cours qui auront un impact sur leur carrière. Ces caractéristiques nous ont amené à considérer l'analyse des guides pédagogiques de sciences et technologie de la 8^e année.

Le guide pédagogique a pour but de proposer aux enseignants des démarches devant leur permettre de mettre en pratique efficacement le curriculum et les aider à suivre le fil du cours. Ainsi, il appuie l'action de l'enseignant en proposant des pistes de travail, des stratégies

d'enseignement et des stratégies d'évaluation (MÉO, 2010). Par ailleurs, les enseignants moins expérimentés se fient beaucoup au guide pédagogique. Le guide pédagogique est un outil qui peut également servir à mettre en place les changements curriculaires et résorber les dérapages perceptibles dans la pratique en offrant à l'enseignant l'information mise à jour tant sur le contenu et la matière que sur les méthodes d'enseignement et d'évaluation nouvellement développées (Beyer, Delgado, Davis et Krajcik, 2009; Lynch, Kuipers, Pyke et Szesze, 2005; Niclot, 2003).

Le rôle du guide pédagogique tel que le suggère l'étude d'Hasni *et al.* (2009) portant sur l'utilisation des manuels scolaires par les enseignants de sciences au regard de la réforme scolaire au Québec, est entre autres d'offrir aux enseignants des activités et des approches d'enseignement facilement utilisables et leur facilitant la préparation des cours. Par ailleurs, l'étude menée par Lin *et al.*, (2010), dont les questions de recherche portaient notamment sur les fonctions et l'aide que les guides pédagogiques pourraient apporter aux enseignants dans le cadre de la mise en place des politiques à Taiwan sous l'influence des réformes scolaires à l'échelle internationale, a montré que pour la plupart des enseignants de sciences le rôle actuel du guide pédagogique est de fournir des savoirs disciplinaires plutôt que des savoir-faire pédagogiques. Selon Lin et ses collègues, les concepteurs de guides pédagogiques du secondaire devraient tenir compte des objectifs de la réforme et rendre les guides pédagogiques plus *pédagogiques*, c'est-à-dire, à même de guider les enseignants dans leur pratique enseignante. Ces deux études portant sur le rôle du guide pédagogique amènent à constater les besoins, soient-ils implicites, des enseignants de sciences en matière d'orientation dans leurs démarches pédagogiques d'une part, et d'autre part, à se rendre compte que ces enseignants en attendent certainement un peu plus afin d'être soutenus dans leurs tâches. Nous pensons que le guide pédagogique serait un outil valable pour aider les enseignants à mettre en œuvre au quotidien des activités d'évaluation formative.

L'évaluation formative est par définition un processus continu et dynamique qui engage l'élève tout au long de sa mise en œuvre et dans lequel certaines activités entreprises tant par l'élève que par l'enseignant fournissent des informations pouvant mener à une rétroaction. Cette rétroaction permet d'apporter des modifications stratégiques sur les apprentissages en cours à travers la régulation de l'enseignement et de l'apprentissage (Black et Wiliam 1998b; Buck et Trauth-Nare, 2009). En outre, dans le processus d'apprentissage des élèves, l'évaluation formative occupe une place de choix parce qu'elle appuie l'enseignant et l'élève dans le coprocessus d'enseignement-apprentissage.

Bien que la fonction première de l'évaluation formative soit d'aider l'élève à ajuster son apprentissage, sa mise en œuvre par l'enseignant n'est pas toujours aussi aisée (Allal, 1989; Bélair, 1989; Black et Wiliam, 1998b; Cizek, 2010; Jalbert et Munn, 2001; Mottier-Lopez, 2007; OCDE, 2005; Scallon, 2000). Elle est assez exigeante pour les enseignants, qui sous la pression des parents plus soucieux des performances, y associent souvent un excès de travail (Halverson, Grigg, Prichett, et Thomas, 2007 ; National Research Council [NRC], 1999). En outre, une évaluation formative devrait être sous-tendue par les principes qui lui sont propres, mais ces principes sont souvent sous-représentés lors de la conceptualisation de l'évaluation formative, raison pour laquelle sa mise en œuvre reste difficile (Bennett, 2011). La mise en œuvre de l'évaluation formative peut donc nécessiter le déploiement de plusieurs moyens, parmi lesquels figure le guide pédagogique (Lynch *et al.*, 2005; Niclot, 2003).

Nous voulons, à travers cette étude, comprendre comment l'analyse des guides pédagogiques de sciences et technologie peut éclairer notre compréhension sur l'existence ou non des principes d'évaluation formative et leur prise en compte. En d'autres termes, nous tentons de répondre à la question générale de recherche suivante : *comment le guide pédagogique de sciences et technologie intègre-t-il les principes d'évaluation formative?* Pour ce faire, nous jetons dans un premier temps un coup d'œil attentif sur la littérature pour voir l'état de la recherche sur le sujet afin de rendre plus précise cette question générale. De cette recension des écrits vont émerger plusieurs principes d'évaluation formative qu'il conviendra, par la suite, de circonscrire dans un cadre conceptuel afin de spécifier l'angle sous lequel nous percevons chacun d'eux. Le cadre conceptuel permettra aussi de spécifier la question de recherche aux contextes entourant notre étude. S'en suivra la méthodologie dans laquelle nous fournissons les informations décrivant l'unité d'étude, les catégories d'analyse, les instruments de collecte de données, la procédure à suivre. Après cela, nous présentons les résultats, et enfin, nous donnons les limites de notre étude ainsi que les contributions qu'elle apportera à la communauté scientifique et à la société.

2.2. Recension des écrits

Comme mentionné dans l'introduction, la première partie de cette recension porte sur le guide pédagogique. Dans la deuxième partie, nous passons en revue les principes de l'évaluation formative, ce qui nous permet de voir comment le guide pédagogique peut contribuer à la mise en œuvre de l'évaluation formative.

2.2.1. Rôle du guide pédagogique

À partir de la définition donnée au guide pédagogique, il est aisé de déduire que celui-ci fait le pont entre le curriculum et les pratiques de salle de classe. On distingue généralement deux catégories de guides pédagogiques : d'une part, on a les guides pédagogiques flexibles et consultatifs, qui laissent à l'enseignant le libre choix de se servir de ses propres principes pédagogiques pour planifier ses leçons et concevoir ses propres ressources pédagogiques (Doyle, 1990). C'est d'ailleurs ce qui est en vigueur en Ontario, où les guides pédagogiques semblent être consultatifs, l'accent étant mis sur les Communautés d'Apprentissage Professionnel (CAP) qui ont pour rôle de favoriser le développement professionnel des enseignants. D'autre part, les guides pédagogiques structurés et prescriptifs procurent à l'enseignant un cadre pédagogique, ainsi que des suggestions et approches pédagogiques, et même probablement des instruments d'évaluation pouvant l'aider à mieux planifier ses unités de cours (Doyle, 1990; Lin *et al.*, 2010).

On peut retenir de ce qui a été dit au sujet du rôle du guide pédagogique, les principaux axes suivants : la nécessité d'une prise en compte de l'apprentissage des élèves dans la réforme, l'utilisation du guide pédagogique par les enseignants et son rôle important dans la prise en main de l'activité pédagogique. Comme le soulignent Reys, Reys et Chávez (2004), le guide pédagogique devrait procurer à l'enseignant de nombreuses activités devant lui permettre d'impliquer les élèves dans leurs apprentissages. Bien sûr, comme le soutiennent Opoku-Amankwa, Brew-Hammond et Kofifah (2011) il n'est pas question que les enseignants s'abandonnent aveuglément au guide pédagogique, mais qu'ils puissent s'en servir afin d'effectuer des évaluations formatives de haute qualité et rendre de ce fait plus prolifiques leurs pratiques d'évaluation formative (McMillan, 2004). Ainsi, nous sommes en accord avec Davis (2006) que la conception de ressources pédagogiques requiert une certaine expertise que tous les enseignants n'ont pas forcément. Comment pourraient-ils alors concevoir des instruments de mesure tout en assurant qu'ils sont pertinents pour mesurer avec fidélité les apprentissages des élèves et produire des résultats valides? À ceci s'ajoute une préoccupation, celle du manque de temps auquel les enseignants sont généralement en proie. De toutes les manières, qu'il soit prescriptif ou consultatif,

un guide pédagogique, dans la mesure où il propose des activités d'évaluation formative, devrait s'assurer que celles-ci respectent les principes qui sous-tendent toute activité d'évaluation formative.

2.2.2. Principes d'évaluation formative

Un principe peut être défini en sciences de l'éducation comme « une proposition générale ou une norme régie par des fondements théoriques et issue soit de la recherche en éducation, soit de l'expérience des enseignants » (Legendre, 2005, p. 1074). En d'autres termes, on pourrait dire que les théories sur lesquelles s'appuie l'évaluation formative sont constituées de principes. Lors de notre recherche documentaire, nous avons trouvé plusieurs écrits portant sur les principes d'évaluation formative. Dans le paragraphe ci-après, nous abordons les auteurs qui ont énuméré des listes de principes dans leurs travaux respectifs.

Les listes de principes d'évaluation formative trouvés dans les écrits sont regroupées sous forme d'un tableau synthèse (Annexe 1). Ce dernier offre une vue d'ensemble des principes et un résumé des explications des auteurs. En 2005, l'OCDE a mené des études de cas à l'échelle internationale dans des pays où l'évaluation formative a fait ses preuves. De l'ensemble de ces études émanent trois principes fondamentaux de l'évaluation formative que les enseignants devraient appliquer afin de créer un environnement d'enseignement et d'apprentissage dynamique et faire progresser les élèves vers les objectifs d'apprentissage préétablis et développer des aptitudes de « savoir apprendre ». Parmi ces principes, figure par exemple l'instauration d'une culture d'évaluation formative de classe qui encourage l'interaction et l'utilisation des outils d'évaluation. Cette culture d'évaluation formative est présentée comme pierre d'angle autour de laquelle viennent se greffer les autres principes. Une autre liste de principes est proposée par Clark (2010) dans un article de recherche théorique sur l'évaluation formative en contexte de réforme. Selon cet auteur, sept principes clés nécessitant une forte et dynamique interaction entre les élèves, les enseignants et les parents sont essentiels à l'établissement d'un lien entre l'enseignement, le curriculum et l'évaluation formative. Le Comité consultatif mixte (1993) a élaboré des principes d'équité relatifs aux pratiques d'évaluation des apprentissages scolaires au Canada en contexte d'évaluation de classe, chacun de ces principes étant suivi d'une série de lignes directrices. Bien que dans l'ensemble ces lignes directrices soient aussi pertinentes les unes que les autres, nous avons résumé dans notre tableau synthèse trois de ces principes portant exclusivement sur l'évaluation formative. Selon Trumpower et Sarwar (2010), l'évaluation formative devrait entre

autres remplir les quatre critères suivants : mesurer les apprentissages pertinents, identifier les forces et les faiblesses de l'élève, fournir une rétroaction efficace et être simple d'utilisation. Pour ces auteurs, l'évaluation formative est étroitement liée à la rétroaction pour l'amélioration des apprentissages. Cizek (2010) propose cinq caractéristiques de l'évaluation formative qui se distinguent des autres en la présence de concept tel que l'autoévaluation. Une dernière série de principes nous est présentée par McMillan (2004) dans un ouvrage portant sur l'évaluation de classe. Nous avons fait une synthèse de ces principes et en avons retenu trois qui portent sur l'évaluation formative. Dans le tableau synthèse, nous rapportons les écrits sur les principes d'évaluation formellement identifiés par leurs auteurs comme tels. Nous y avons inclus les principes qui, selon nous, portent sur les activités d'évaluation formative et en avons exclu ceux qui portent uniquement sur les activités d'enseignement ou d'évaluation sommative.

En dehors des travaux des précédents auteurs, certains auteurs clés du domaine de l'évaluation formative se sont appuyés sur les principes d'évaluation formative dans leurs travaux. Bien qu'ils n'aient pas, à notre connaissance, dressé une liste formelle de principes, les principes qu'ils mentionnent ont également le mérite d'être pris en compte dans le cadre de cette étude compte tenu de la renommée de ces auteurs et leur influence dans le domaine. Par exemple, Scallon (1988a) préconise que les scénarios d'évaluation formative soient souples. Et dans cette foulée, ce même auteur en 1999 propose quatre caractéristiques de souplesse d'une évaluation formative :

1. il n'y a pas toujours de résultat chiffré;
2. ce n'est pas toujours l'élève qui est jugé. À notre sens, l'objet d'évaluation ne porte pas sur l'élève, mais sur les habiletés, voire la pédagogie « (entendue ici dans un sens très général) » (p. 8);
3. on vérifie des apprentissages inachevés;
4. l'élève participe à son évaluation. « L'évaluation formative doit emprunter des mises en scène multiples et tout aussi souples les unes que les autres: autocorrection, co-évaluation, travail coopératif, exercices hors classe » (p. 8).

Par ailleurs, la recension portait à la fois sur le rôle des guides pédagogiques et sur les principes d'évaluation formative des apprentissages. Compte tenu de la question générale de recherche à savoir comment le guide pédagogique contribue à la mise en place de l'évaluation formative, nous faisons, dans la prochaine section, le lien entre le guide pédagogique et

l'évaluation formative en mettant en symbiose les principes qui émergent de la recension et les concepts qui sont définis dans ladite section, afin de rejoindre cette question de recherche.

2.3. Cadre conceptuel

Le cadre conceptuel de la présente étude permet de décrire et situer en un tout cohérent, les concepts clés qui émergent de notre recension des écrits. Dans cette section, nous analysons ces concepts clés et les mettons en relation les uns avec les autres par regroupement thématique. En effet, le regroupement des principes d'évaluation formative retrouvés dans la littérature a permis de constater l'émergence de deux dimensions : le processus et les conditions. Premièrement et comme le souligne Popham (2008), l'évaluation formative est un processus et non un test, contrairement à ce qui est communément accepté. De nombreux principes sont vraisemblablement liés au processus d'évaluation. Deuxièmement, nous remarquons la présence de principes touchant aux conditions de mise en œuvre de l'ensemble du processus d'évaluation formative. Nous décrivons en profondeur ces deux dimensions ainsi que les concepts associés aux principes d'évaluation formative dans les prochaines sous-sections.

2.2.1. Processus

Le processus d'évaluation formative, comme toute démarche évaluative, est constitué d'étapes. Forgette-Giroux et Simon (1998) en proposent quatre :

1. intention : l'intention est une étape de planification qui consiste à établir : la fonction d'évaluation – par exemple, une évaluation pourrait avoir pour fonction la régulation des apprentissages ou le diagnostic –; le responsable d'évaluation qui peut être l'évalué, l'enseignant ou le pair; et l'objet d'évaluation qui indique précisément sur quoi porte l'évaluation;
2. mesure : lors de cette étape, il est question de déterminer la tâche et l'instrument de cueillette d'informations appropriées qui permettront de mesurer, observer ou capter le niveau de rendement démontré par l'élève;
3. jugement : le jugement permet de situer le rendement que l'élève a démontré à l'étape de mesure, de manière objective et uniforme selon des critères résultant de l'opérationnalisation de l'objet d'évaluation, et par l'usage d'outils d'évaluation;
4. décision et communication : elles consistent à choisir entre différentes actions permettant de réajuster la démarche d'apprentissage de l'élève à travers la régulation de

l'enseignement (Forgette-Giroux et Simon, 1998). Les nombreux travaux autour de la rétroaction font ressortir certaines caractéristiques de la rétroaction : clarté, descriptivité, précision, opportunité, ponctualité, fréquence, signifiante, etc. (Allal et Mottier Lopez, 2005 ; Brookhart, 2010 ; Cizek, 2010 ; Earl et Katz, 2006 ; Hattie et Timperley, 2007 ; OCDE, 2005 ; Scallon, 1988b).

2.2.2. Conditions

Les conditions entourant les activités d'évaluation formative sont les suivantes :

1. la souplesse et le dynamisme : la souplesse réfère à l'application du processus d'évaluation, et consiste par exemple, en l'autorisation des changements aux modalités de l'activité d'évaluation en cours. Pour le dynamisme, il est question de rendre fructueux les échanges élève – enseignant à travers une discussion, comme cela se ferait lors de la vérification de la compréhension des critères par un élève au début de l'évaluation;
2. le droit à l'erreur : les erreurs ne devraient aucunement être utilisées pour sanctionner la compétence de l'élève, mais comme une occasion d'autorétroaction significative pour se réajuster par la suite;
3. la différenciation : c'est l'adaptation des méthodes d'évaluation au bagage de chaque élève et la promotion de l'égalité de chances pour l'amélioration des apprentissages de chaque élève et de tous les élèves (Laveault, 2010);
4. la responsabilisation : selon Allal (1993), elle découle d'une réflexion métacognitive devant enclencher chez l'élève une autorégulation de ses apprentissages. Elle permet également à l'élève de participer à toutes les étapes du processus d'évaluation formative et l'incite à faire appel à ses pairs, etc. La responsabilisation englobe les processus métacognitifs et autoévaluatifs;
5. l'interactivité : il s'agit d'une interaction entre les élèves dans le sens de faire participer ceux-ci à l'évaluation de leurs pairs (Laurier, Tousignant et Morissette, 2005; Wiggins, 1993).

Nous venons de définir un cadre constitué de concepts clés qui ont été développés à la lumière des écrits scientifiques. Ces concepts tels que définis servent de cadre aux principes d'évaluation formative énumérés dans la recension des écrits, aux fins de cette étude. Les propos que nous avons évoqués depuis le début permettent par ailleurs de constater le besoin, sur le plan de la recherche, de mettre en lumière la contribution du guide pédagogique au succès de

l'apprentissage de l'élève à travers les activités qu'il propose aux enseignants. Par conséquent, il serait d'un certain intérêt de voir comment les principes d'évaluation formative éventuellement présents dans les guides pédagogiques participent à rendre formatives les activités d'évaluation qu'on y retrouve dans le but d'améliorer l'enseignement. Ce qui nous amène à formuler la question spécifique de recherche comme suit : *De quelle manière les activités d'évaluation proposées dans les guides pédagogiques de sciences et technologie de l'Ontario à des fins formatives prennent-elles en compte des principes d'évaluation formative?* Pour tenter de répondre à cette question spécifique de recherche, nous avons eu recours à un cadre méthodologique adapté.

2.4. Méthodologie

L'analyse de contenu de guides pédagogiques est une opération qui permet d'effectuer une analyse rigoureuse des textes et des messages contenus dans les documents, en utilisant des techniques scientifiques, objectives et systématiques d'analyse de textes (Richard, 2006; Gohier, 2011). Cette section montre comment s'est opérationnalisée l'étude et s'étale sur quatre parties : les unités d'étude et unités d'analyse, les catégories analytiques, les instruments et la procédure.

2.2.1. Unités d'étude et unités d'analyse

Au moment de l'élaboration de cette étude, quatre guides pédagogiques de sciences et technologie de 8^e année figurent dans la Liste Trillium : deux provenant du secteur francophone et deux autres du secteur anglophone. Les deux guides pédagogiques francophones sont retenus parce qu'ils sont des traductions des guides pédagogiques anglophones et ceci permet de ne pas collecter de doublons. Il s'agit des guides pédagogiques Perspectives Sciences et technologie 8 de Duval publié en 2009 et Investigation 8 Sciences et technologie d'ERPI publié en 2010 ci-après abrégés GPP et GPI respectivement. En plus de ces deux guides pédagogiques approuvés, le guide pédagogique Technoscience, ci-après abrégé GPT, publié en 2001, est selon le Centre Franco-Ontarien des Ressources Pédagogiques (CFORP) parmi les plus populaires chez les enseignants (CFORP, communication personnelle, 21 mars 2013). Pour cette raison, nous l'avons ajouté à notre corpus. Bien que ces guides pédagogiques aient été édités avant ou pendant l'année de publication de la politique d'évaluation de l'Ontario, nous pensons que les principes d'évaluation formative devraient quand même y être perceptibles parce que le GPP et le GPI sont les seules ressources approuvées par le MÉO. De fait, la réforme en Ontario est passée par plusieurs étapes jusqu'à sa consolidation et donc le reflet pourrait être perceptible dans des guides pédagogiques

édités en 2009 et 2010. De plus, l'usage actuel du GPT en fait un sujet d'intérêt. Nous nous retrouvons donc avec un corpus de trois guides pédagogiques comme sources de données : le GPP, le GPI et le GPT. Étant donné que nous sommes intéressé par la technologie, nous retenons le domaine intitulé Structures et mécanismes de chaque guide pédagogique parce que ces domaines rejoignent les concepts technologiques du curriculum de l'Ontario de 9^e année et 10^e année en éducation technologique. Le domaine des structures et mécanismes est un des quatre domaines couverts par le guide pédagogique. Ceci revient à dire que nous analysons une unité du guide sur quatre.

Les unités d'analyse sont les activités et les rubriques qui sont formellement identifiées dans les guides pédagogiques comme activités d'évaluation formative. Les rubriques tiennent le plus souvent sur une demi-page. Les activités s'étendent généralement sur plusieurs pages, sont plus complètes et détaillées. Le terme « activité » est utilisé par la suite pour désigner ces activités, rubriques et tâches. En somme, on dénombre 40 activités d'évaluation formative dans le GPP; 21 dans le GPI et deux dans le GPT. Chaque guide pédagogique a un contenu et un préambule. Les données contenues dans le préambule concernent généralement l'ensemble du guide pédagogique et les données provenant du contenu sont propres à chaque activité.

2.2.2. Catégories analytiques et système de codage

Les catégories analytiques ont été arrimées à chacune des deux dimensions retenues dans le cadre conceptuel, et les codes arrimés aux concepts clés. Ainsi, huit codes ont été générés pour le processus et cinq codes pour les conditions. Malgré la prédétermination des codes, nous sommes resté ouvert à l'émergence d'autres codes durant la collecte de données. Puisque les catégories analytiques découlent des deux dimensions et étant donné la structure de chaque dimension – par exemple les étapes consécutives du processus – et la disparité des deux dimensions – par exemple, le processus est de nature séquentielle et les conditions sont de nature combinatoire –, deux instruments de collecte de données distincts ont été conçus : l'instrument de collecte de données relatives au processus (Tableau 3) et l'instrument de collecte de données relatives aux conditions (Tableau 4)

La grille d'analyse des données provenant de chaque instrument comprend les unités d'analyse et les informations recueillies issues de la collecte de données. L'analyse des données porte d'une part sur l'occurrence des principes; par exemple, on cherche à savoir si le principe est

présent ou absent. Et d'autre part, sur la qualité de la prise en compte du principe; pour ce faire, on se sert des attributs suivants : complète, suffisante, cohérente, pertinente, à jour, désuète, explicite, représentative, etc.

2.2.3. Instruments

Les instruments de collecte de données sont les suivants : l'instrument de collecte de données relatives au processus (Tableau 3) et l'instrument de collecte de données relatives aux conditions (Tableau 4)¹. Ces instruments reprennent la synthèse des principes de l'Annexe 2. Ces principes sont transformés en questions guides de collecte de données, auxquelles nous avons essayé lors de la collecte de données de trouver des réponses, en puisant dans les guides pédagogiques des extraits aux fins d'analyse. Les instruments ont été éprouvés par des exercices de validité.

La première étape de validité a été de conformer les instruments aux règles d'élaboration des listes de vérification. Ensuite, nous avons éliminé les questions redondantes et reclassé les questions retenues pour les faire correspondre à chacun des principes préalablement synthétisés. Ceci a permis de rendre les codes intelligibles et mutuellement exclusifs.

Ensuite, nous avons fait un exercice d'appariement en demandant à une collègue en mesure et évaluation d'associer les questions guides des instruments de collecte de données aux codes préalablement dispersés. Tous les codes ont été appariés correctement. Cet exercice a ainsi permis de valider le contenu des instruments de collecte de données.

Par ailleurs, nous avons procédé à une étude pilote qui a permis de vérifier l'exclusivité mutuelle des codes et d'établir un indice de fidélité interjuge. À cette étude pilote s'est ajouté un chercheur autre que l'auteur, et ce dans le but d'augmenter la fiabilité et la crédibilité de l'étude (L'Écuyer, 1987). Ce deuxième chercheur est un collègue qui a reçu une formation de manière à pouvoir effectuer la collecte de données de façon autonome. Après avoir balayé à l'aide des instruments certaines sections d'un guide pédagogique choisi au hasard, nous avons colligé nos résultats et les avons comparés. Lors de cette phase de validation d'instrument de collecte de données, chaque code renseigne sur la présence ou l'absence des principes d'évaluation formative. Ceci étant, les accords sur les occurrences et sur les non-occurrences des principes sont nécessaires.

¹ Le Tableau 3 et le Tableau 4 seront annexés à l'article avant publication.

À cet effet, Hawkins et Dotson (1975, cité par House, House et Campbell, 1981) proposent la formule suivante pour calculer la moyenne (M) du pourcentage d'accords sur les occurrences et les nonoccurrences : $M(\%) = (A/(A + B + C) + D / (B + C + D)) / 2 \times 100$. Nous avons obtenu une moyenne de 77%. Ce qui selon House, House et Campbell est acceptable.

2.2.4. Procédure

La collecte de données a immédiatement suivi l'étude pilote. Lors de l'examen des préambules, la structure de chaque document a été étudiée pour voir comment sont conçues l'évaluation formative et les activités d'évaluation identifiées comme formatives dans les unités d'étude. Toutes les données relatives aux principes d'évaluation formative pouvant être mentionnées explicitement ou implicitement dans le préambule ont été extirpées. C'est ainsi qu'on pouvait par exemple ressortir des constatations relatives à la rétroaction. Ensuite, à l'aide de chacun des deux instruments de collecte de données, les contenus des guides pédagogiques ont été à leur tour balayés; ce faisant, on prenait soin de noter dans les instruments tout extrait pertinent. Pendant la collecte de données, nous sommes resté ouvert à l'émergence de codes. Cependant, nous n'avons pas eu de nouveaux codes. Les analyses sont de nature qualitative dans la mesure où chacun des extraits nous renseigne sur la manière dont est pris en compte le principe d'évaluation formative.

2.5. Résultats, interprétation et discussion

Les résultats de l'étude sont présentés dans cette section et de différentes façons. Premièrement, il y a les résultats qualitatifs qui consistent en la description des données collectées pour chaque code. Ces résultats sont regroupés en fonction des codes ayant servi à la collecte de données. À la suite de chaque résultat, nous donnons une interprétation d'une part, et d'autre part, nous discutons sur la façon dont ce résultat permet de se prononcer sur la qualité de la prise en compte du principe. Lorsque la discussion porte sur plus d'un concept, celle-ci est faite après l'interprétation du résultat du tout dernier concept. Deuxièmement, nous avons associé à chaque dimension un tableau pour quantifier les données de façon à ce que le lecteur puisse avoir une vue plus globale de la prise en compte des principes. Il s'agit du Tableau 5 et du Tableau 6. Troisièmement, ces mêmes données sont présentées par guides pédagogiques, et ce de façon plus détaillée dans l'Annexe 3, l'Annexe 4 et l'Annexe 5. À titre d'illustration, nous en avons inséré dans le texte l'Annexe 4, c'est-à-dire l'occurrence des données provenant du GPI (Tableau 7). Soulignons que nous ne prétendons pas faire une analyse quantitative de ces données parce que

leur nombre ne justifierait pas un tel emploi. Mais nous essayons de les quantifier dans le seul but d'avoir un portrait de l'occurrence des principes dans les activités. Enfin, nous avons produit en Annexe 6 un tableau dans lequel figure un exemple d'extrait relatif à chaque code lorsque présent dans les unités d'étude.

Processus – Intention – Fonction d'évaluation

La fonction d'évaluation est indiquée une fois dans le GPP; on peut lire dans l'activité de fin d'unité :

« L'activité de fin d'unité vise deux objectifs : l'évaluation pour l'apprentissage et l'évaluation de l'apprentissage. Cette tâche est un point culminant permettant aux élèves de consolider leur apprentissage. Elle leur permet également de manifester leur compréhension des concepts et de mettre en pratique les habiletés acquises au cours de l'unité. Ils peuvent ainsi montrer qu'ils saisissent bien comment les sciences et la technologie abordées dans cette unité influencent la société et l'environnement » (GPP, 2010, p. 115).

Dans l'extrait précédent, les fonctions d'évaluation semblent naturellement découler des objectifs d'apprentissage indiqués. Étant donné que le préambule du GPP associe l'évaluation pour l'apprentissage à l'évaluation formative, on peut considérer que la fonction d'évaluation est bel et bien indiquée au début de cette activité. De surcroît, « elle permet aux élèves de consolider leur apprentissage », ce qui est tout à fait propre à l'évaluation formative. Cependant, il n'y a pas d'autres indications de la fonction d'évaluation dans ce guide pédagogique. Dans le GPI et le GPT la fonction d'évaluation n'est pas indiquée.

Processus – Intention – Responsable de l'évaluation

Dans le GPP, le responsable d'évaluation est spécifié dans 19 rubriques et activités; par exemple : « Vous pouvez recueillir le DR 1.5-1 et évaluer les réponses des élèves au moyen de la grille d'évaluation 7, 'Prononce-toi sur un enjeu' » (GPP, 2010, p. 40). Dans cet extrait comme dans plusieurs autres extraits issus de la collecte de données, on n'indique pas explicitement qui est le responsable d'évaluation. On peut cependant inférer de l'ensemble de l'extrait que le responsable est l'enseignant. Le responsable d'évaluation est donc implicitement identifié dans cet extrait. À quelques reprises cependant, le responsable d'évaluation est explicitement indiqué; par exemple : « les élèves peuvent évaluer leurs habiletés à **élaborer** une démarche expérimentale (...) au moyen de la liste de vérification de l'autoévaluation 5 (...) » (GPP, 2010, p. 74). Dans ce deuxième extrait, les responsables d'évaluation sont les élèves. Dans 7 des 19 extraits issus de la collecte de données, le responsable d'évaluation est l'élève. Il est important que l'élève soit

responsable de ses apprentissages en effectuant des activités d'autoévaluation. C'est pourquoi en contexte d'évaluation formative on favorise les activités autoévaluées (Clark, 2010). Il n'y a pas de données portant sur le responsable d'évaluation dans le GPI et le GPT. Soulignons que dans les deux extraits précédents, il y a certaines incohérences au niveau de l'objet d'évaluation qui seront explicitées dans la section suivante.

Processus – Intention – Objet d'évaluation

L'objet d'évaluation est mentionné dans 17 activités du GPP, dans une activité du GPI, et dans aucune activité du GPT. Voici un exemple d'objet d'évaluation en bonne et due forme : « Cette activité permet d'évaluer la **compréhension** des élèves au sujet du travail » (GPI, 2009, p. 15). Dans cet exemple, le responsable d'évaluation évalue la compréhension de la notion du travail. Cependant, quatre problèmes sont à relever au sujet de l'objet d'évaluation ici, à savoir : sa nature, sa pertinence, sa formulation et son doublage.

Premièrement, reprenons les deux extraits de la section précédente :

« Vous pouvez recueillir le DR 1.5-1 et évaluer les réponses des élèves au moyen de la grille d'évaluation 7, 'Prononce-toi sur un enjeu' » (GPP, 2010, p. 40).

« Les élèves peuvent évaluer leurs habiletés à **élaborer** une démarche expérimentale (...) au moyen de la liste de vérification de l'autoévaluation 5 (...) » (GPP, 2010, p. 74).

En regardant ces extraits, on peut remarquer dans le premier qu'on demande au responsable d'évaluation d'évaluer des réponses, et dans le second, le responsable d'évaluation est appelé à évaluer des habiletés. L'évaluation d'habiletés nous semble plausible dans la mesure où ces habiletés sont indiquées dans les échelles d'appréciation. Cependant, lorsqu'on demande d'évaluer des réponses, on ne sait pas vraiment quel est l'objet d'évaluation parce qu'il n'y a pas matière à évaluer. On retrouve à plusieurs reprises ce genre d'incohérences sur la nature de l'objet d'évaluation dans les activités d'évaluation formative dans le GPP.

Deuxièmement, il semble avoir dans le GPP, des problèmes de pertinence se traduisant par un manque d'alignement entre l'objet d'évaluation et les critères d'évaluation (De Ketele et Gérard, 2005). Citons-en un exemple :

« Vous pouvez également leur demander de former des équipes pour la préparation d'une présentation. Vous évalueriez la qualité de leurs explications et de leurs idées au moyen de la Grille d'évaluation 1, «Connaissance et compréhension», et de la Grille d'évaluation 3, «Communication» (GPP, 2010, p. 40).

Dans cet exemple, on demande d'évaluer la qualité des explications et des idées des élèves en servant des échelles d'appréciation désignées (grille d'évaluation 1 et 3). Les critères d'évaluation spécifiés dans ces échelles sont les suivants :

« Connaissance des éléments à l'étude (p. ex., définir des termes à l'étude, concepts, lois, théories, consignes de sécurité);
Compréhension des éléments à l'étude (p. ex., reconnaître des concepts, principes, lois, théories);
Expression et organisation des idées et de l'information » (GPP, 2010, p. OÉ-1).

« Communication des idées et de l'information, de façon orale, écrite et visuelle à des fins précises (p. ex., présenter, informer, défendre) et pour des auditoires spécifiques (p. ex., camarades, personnel enseignant, communauté);
Utilisation des conventions (p. ex., symboles, formules, unités SI) et de la terminologie à l'étude » (GPP, 2010, p. OÉ-3).

Il appert qu'aucun de ces critères d'évaluation n'est en lien avec l'objet d'évaluation, en l'occurrence l'évaluation de la qualité des explications et des idées des élèves. Ce qui pose un problème de pertinence parce que l'évaluation ne touche pas les cibles originelles (De Ketele et Gérard, 2005).

Troisièmement, la manière de formuler l'objet d'évaluation peut faire en sorte que celui-ci passe inaperçu. Les exemples 1 et 2 ci-après tirés du GPP et du GPI illustrent ce propos.

Exemple 1 :

« Habiletés : Exécuter, observer, analyser, évaluer et communiquer » (GPP, 2010, p. 126).

« Habiletés : analyser et interpréter, communiquer » (GPI, 2009, p. 15).

Exemple 2 :

« Habiletés : Les sections *En résumé* et *Révision* permettent aux élèves de montrer leur **compréhension** et leur **aptitude à utiliser** les concepts clés, le vocabulaire et les habiletés » (GPP, 2010, p. 111).

« Cette activité permet d'évaluer la **compréhension** des élèves au sujet du travail » (GPI, 2009, p. 15).

Dans l'exemple 1, on voit une liste de verbes cités qui semblent faire office d'objets d'évaluation, ce d'autant plus que lorsque l'on regarde de près l'échelle d'appréciation proposée qui est la *Grille d'évaluation 3 : Communication* pour le cas du GPP, on retrouve un lien entre certaines habiletés de l'exemple 1 et les critères d'évaluation contenus dans l'échelle d'appréciation. Dans l'exemple 2, l'objet d'évaluation est développé de façon à faire ressortir l'apprentissage et la matière à évaluer, que nous avons respectivement mis en gras et en souligné dans les exemples et

les extraits précédents. Dans l'éventualité où l'on considérerait les habiletés énumérées dans l'exemple 1 comme des objets d'évaluation, il y aurait une incohérence entre les formulations des objets d'évaluation présentés dans l'exemple 1 et l'exemple 2. Le cas échéant, l'objet d'évaluation passerait inaperçu dans l'exemple 1. En outre, l'exemple 1 comporte un problème de pertinence que l'on retrouve à quelques autres reprises dans le GPP, et rejoint notre remarque au sujet de l'alignement entre les objets d'évaluation et les critères d'évaluation. Par exemple, les habiletés comme l'observation et l'analyse n'ont pas de critères d'évaluation et par conséquent, l'évaluation manque de pertinence (De Ketele et Gérard, 2005).

Quatrièmement, les deuxièmes extraits de l'exemple 1 et de l'exemple 2 sont tirés de la même activité dans le GPI. Dans un extrait, on évalue la compréhension des élèves au sujet du travail, et dans l'autre on fait une analyse, une interprétation et une communication. Ceci étant, on fait face à un double objet d'évaluation pour la même activité. Ce double objet d'évaluation complexifie l'activité d'évaluation formative qui est censé être simple d'utilisation selon Trumpower et Sarwar (2010).

En guise de discussion en lien avec cette étape d'intention relative à la fonction d'évaluation, au responsable d'évaluation et à l'objet d'évaluation, nous pouvons souligner à l'instar de Forgette-Giroux et Simon (1998) que l'intention est une étape de planification relevant du processus et à cet égard, tout le processus d'évaluation pourrait être ébranlé si cette étape était bâclée. Or, la fonction d'évaluation, comme le montre le Tableau 5 portant sur la quantification des données relatives au processus, n'est presque pas prise en compte dans les guides pédagogiques. Pourtant, à notre sens la fonction prépare l'évalué à ce qui l'attend et l'aide à se disposer en conséquence. De fait, dire à un élève qu'on va l'évaluer sommativement au cours d'une activité n'a pas le même effet que si on lui annonce qu'il le sera formativement. En effet, les mécanismes propres à l'évaluation formative dont nous avons parlé dans le cadre conceptuel tels que le droit à l'erreur peuvent s'activer instantanément lorsqu'un élève sait qu'il est en posture formative, et peut dès lors oser prendre des risques dans sa démarche de construction de savoirs, et apprendre beaucoup mieux. Si l'on considère qu'au moment de notre étude, nous avons sélectionné les activités à analyser sur la base de l'indication explicite de la fonction d'évaluation dans les préambules, on pourrait considérer que les guides pédagogiques indiquent la fonction d'évaluation dans les préambules. Mais ceci semble insuffisant parce que lorsque nous prenons en compte le temps qu'il nous a fallu consacrer à éplucher ces préambules afin de dénicher les

activités d'évaluation formative, on se dit que la tâche de l'enseignant aurait été simplifiée si la fonction d'évaluation était explicitement indiquée dans les activités. En ce qui concerne le responsable d'évaluation, les résultats montrent qu'on met le plus souvent l'accent sur l'enseignant plutôt que l'élève, malgré le fait qu'on soit dans un contexte d'évaluation formative. Or, dans un tel contexte, le responsable d'évaluation passe souvent de l'enseignant à l'élève ou aux pairs. Par conséquent, la désignation du responsable d'évaluation revêt une certaine importance. Nous soutenons que le responsable d'évaluation devrait être explicitement indiqué par souci de transparence également, ce afin d'éviter toute équivoque. Pour ce qui est de l'objet d'évaluation, il y a lieu de se questionner sur la pertinence des activités d'évaluation formative et la validité de leurs résultats quand on découvre qu'aucune des activités d'évaluation formative, dans un guide donné ou bien que la plupart des activités dans un autre guide, n'ont pas d'objet d'évaluation. Et quand bien même il y en a un, des problèmes de pertinence et de cohérence altèrent son effet et influencent sa prise en compte dans les activités d'évaluation formative. Nous croyons par ailleurs que l'indication dans le préambule des intentions d'évaluation, c'est-à-dire la fonction, le responsable et l'objet d'évaluation ne suffit pas. Il s'agit également de communiquer à l'élève ces informations en début d'activité afin de le disposer à entamer l'activité dans un esprit formatif.

Tableau 5 : Quantification des données par dimension – Processus

Processus	GPP <i>n</i> = 40	GPI <i>n</i> = 21	GPT <i>n</i> = 2
Intention – Fonction d'évaluation	1	0	0
Intention – Responsable d'évaluation	19	0	0
Intention – Objet d'évaluation	17	1	0
Mesure – Tâches	23	13	2
Mesure – Instruments	9	0	0
Jugement – Outils de jugement	11	0	2
Jugement – Critères d'évaluation	11	0	2
Communication – Rétroaction	0	0	0

n représente le nombre d'unités d'analyse de chaque guide pédagogique

Le GPP présente 23 tâches et 9 instruments. En revanche, certains instruments et tâches sont soit incomplets, soit inaccessibles. Par exemple, les activités *Vérifie ta compréhension* n'ont ni tâches, ni instruments. Cette activité est constituée de réponses à des questions qui sembleraient se trouver dans un document connexe, en l'occurrence le manuel de l'élève. Le GPI comporte 13 tâches et aucun instrument de mesure. Ces tâches semblent pour la plupart incomplètes. En guise d'exemple, nous citons intégralement une activité dite formative :

« B12 Fais le point ! : Calculer le travail

- Dire aux élèves de relire leurs notes et la section *Calculer* le travail avant de répondre à ces questions.
- Permettre aux élèves de se servir d'une calculatrice pour répondre aux questions. Leur rappeler de commencer par écrire la formule, puis de faire les substitutions à l'aide des unités appropriées.

Réponses aux questions

1. 30 J
2. 8,0 J » (GPI, 2009, p. 19).

Dans l'extrait précédent, on voit des instructions relatives au déroulement de la tâche d'évaluation formative. Mais, il n'y a pas de tâche d'évaluation à proprement parler, ni d'instruments de mesure. Pour le GPT, la tâche d'évaluation est présente dans les deux activités d'évaluation formative proposées comme le montre l'extrait suivant : « Proposer aux élèves de faire un dessin technique, de fabriquer le prototype amélioré d'un fauteuil roulant ou de faire tout autre véhicule semblable, lui permettant de surmonter les petits obstacles (...) » (GPT, 2001, p. 31). Une tâche d'observation comme celle-ci demande que le responsable d'évaluation observe l'élève pendant les travaux de réflexion de l'équipe et, sur un instrument de mesure comme une grille d'observation (par exemple), consigne les constatations en se référant aux critères d'évaluation. Remarquons que cette tâche d'évaluation est en lien avec la technologie au sens du MÉO, c'est-à-dire basée sur la réalisation d'objet technologique. Aucun instrument de mesure n'est présent dans le GPT.

La moitié des activités proposées dans la plupart des guides pédagogiques sont pourvues de tâches. Ceci semble aberrant lorsqu'on se pose la question de savoir comment s'effectueraient alors l'activité d'évaluation formative avec seulement la moitié des tâches. Et quand bien même il y aurait des tâches, celles-ci ne sont pas toujours complètes. Une tâche absente ou incomplète est difficilement utilisable par l'enseignant, qui devrait alors se rabattre dans les ressources connexes afin de trouver les parties manquantes. La présence de la tâche dans le guide aurait permis à l'enseignant d'avoir accès à toute l'information sans avoir à feuilleter le manuel de l'élève à tous

les coups. En ce qui concerne les instruments, on en trouve uniquement dans le GPP et dans le tiers des activités proposées. Pourtant, ces instruments devraient être non seulement présents et variés en ce qui concerne les niveaux taxonomiques des habiletés à évaluer comme le suggère Gronlund (2003), mais aussi diversifiés en matière de stratégie d'évaluation – comme des productions ou des observations – tel que le stipule l'OCDE (2005), tout cela afin de répondre aux styles d'apprentissage diversifiés des élèves. Par ailleurs, selon McMillan (2004) et Trumpower et Sarwar (2010), le processus d'évaluation formative devrait permettre de mesurer les apprentissages pertinents. Comment ceci peut-il se faire lorsque dans certains guides pédagogiques, les instruments et les tâches sont absents, incomplètes ou inaccessibles? Les apprentissages pertinents au sens de Trumpower et Sarwar portent sur les habiletés supérieures à la connaissance et compréhension, comme la pensée critique ou la mise en application. Ces habiletés devraient être mesurées selon McMillan par des instruments comme le dossier d'apprentissage, par exemple, qu'on ne retrouve dans aucun guide pédagogique à l'étude.

Processus – Jugement – Outils de jugement (outil d'évaluation)

Dans le préambule du GPP, il est écrit que : « Les *Outils d'évaluation* contenus dans le présent *Guide d'enseignement* comprennent les *Grilles d'évaluation*, les *Résumés de l'évaluation* (visant le rendement individuel de l'élève) et les *Listes de vérification de l'autoévaluation* » (GPP, 2010, p. xvii). Cet extrait suggère que le GPP regorge d'outils d'évaluation, et précise également le type d'outil d'évaluation que l'on peut trouver dans cette ressource. Lorsque l'on parcourt le contenu du GPP, on retrouve en effet des outils d'évaluation dans 11 activités; par exemple : « Vous pouvez circuler et discuter avec les élèves pendant l'exercice de résolution de problèmes. Évaluez leurs habiletés à communiquer et à résoudre des problèmes au moyen de la Grille d'évaluation 1, 'Connaissance et compréhension' » (GPP, 2010, p. 60).

Le préambule du GPI suggère l'usage d'outils de jugement tels que des échelles d'appréciation (grilles d'évaluation), des règles de notation (guide de notation) et des critères d'évaluation. Cependant, on ne les trouve pas dans le guide pédagogique. Encore ici, on a affaire à un problème d'accessibilité de ressources.

Les outils de jugement sont présents dans le GPT et constitués de deux échelles d'appréciation : la grille B et la grille C, contenant des critères d'évaluation.

Processus – Jugement – Critères d'évaluation (critères de réussite)

Les critères d'évaluation sont énumérés dans l'intégralité des échelles d'appréciation génériques que propose le GPP. Par exemple, dans l'extrait qui suit, il est demandé à l'enseignant de se servir de deux échelles d'appréciation pour évaluer la qualité des explications et des idées des élèves : « Vous évalueriez la qualité de leurs explications et de leurs idées au moyen de la Grille d'évaluation 1, «Connaissance et compréhension», et de la Grille d'évaluation 3, «Communication» » (GPP, 2010, p. 40); et lorsqu'on examine ces échelles d'appréciation, on constate la présence des critères d'évaluation comme la « Connaissance des éléments à l'étude [et la] Compréhension des éléments à l'étude » (GPP, 2010, p. OÉ-1). Cependant, ces critères ne sont pas toujours alignés avec l'objet d'évaluation comme nous l'avons souligné dans la section portant sur l'objet d'évaluation. Par exemple, l'objet d'évaluation requiert l'évaluation de la qualité des explications des élèves. Mais dans l'échelle d'appréciation préconisée (grille d'évaluation 1), on propose l'usage du critère connaissance des éléments à l'étude et non la qualité des explications ou la qualité des idées. Ces échelles ne sont donc pas adaptées aux objets d'évaluation et il y a un manque de pertinence (Laurier et *al.*, 2005). Par conséquent, on ne sait plus ce qu'on évalue et en plus toutes les parties prenantes de l'activité d'évaluation pourraient avoir des interprétations différentes des critères d'évaluation.

Dans le GPI, il est recommandé dans le préambule que les enseignants se servent des copies types pour comprendre les critères d'évaluation. On entend ici par copie type des exemples de travaux d'élèves pour lesquels des notes couvrant l'ensemble de l'échelle de notation (niveaux 1 à 4) ont été attribuées, ce de façon à guider les responsables d'évaluation lors de l'étape de notation et appréciation. Donc, la seule mention des critères d'évaluation fait allusion aux copies types. Or, en Ontario, l'usage des copies types a été abandonné il y a quelques années. De plus, de prime abord les copies-types de même que les grilles d'évaluation génériques ont été conçues pour rendre la notation des élèves plus rigoureuse. Leur usage à des fins formatives, quoique possible, particulièrement au niveau de la transmission des objectifs et des critères aux élèves, est discutable. Ce qui revient à dire qu'aujourd'hui, les enseignants n'ont aucun repère pour comprendre les critères d'évaluation tels que spécifiés dans le préambule du GPI, ni par ricochet, les faire comprendre par l'élève. Par ailleurs, il ne s'agit pas seulement de comprendre les critères d'évaluation, mais aussi et surtout d'y avoir accès; ce qui n'est pas le cas. En somme ici, on peut relever un problème de désuétude de la ressource proposée par un guide pédagogique encore en

usage; et également des problèmes d'occasions incomplètes, de ressources inaccessibles ou non pertinentes.

Pour le GPT, les critères d'évaluation sont contenus dans les deux échelles d'appréciation proposées; par exemple : « Nous avons bien échangé avec les autres équipes l'analyse des dessins techniques (...) Nous avons suivi les étapes du processus de design pour fabriquer le prototype ». Ces critères sont suivis de règles de notation. Un problème ici c'est l'absence d'objet d'évaluation, qui rend difficile une quelconque interprétation de cette constatation. De plus, il semble ne pas avoir de lien entre les critères d'évaluation et la tâche; par exemple, dans la *Grille d'évaluation par les pairs* (grille C) associée à la tâche de la page 31 du GPT que nous avons citée plus haut, il y a entre autres critères, les suivants : « Nous avons bien compris les exigences; nous avons trouvé plusieurs applications courantes de la pneumatique, etc. » (GPT, 2001, p. 42). Ces critères d'évaluation n'ont pas de lien avec la tâche d'évaluation, laquelle propose aux élèves de faire un dessin technique, de fabriquer le prototype amélioré d'un fauteuil roulant ou de tout autre véhicule semblable, lui permettant de surmonter les petits obstacles. On voit bien qu'aucune allusion n'est faite des applications courantes de la pneumatique, par exemple. Ce qui pose un problème de pertinence. Remarquons cependant que dans l'extrait précédent, la notion de pneumatique est en lien avec la technologie.

En somme pour ce qui a trait à cette étape de jugement, on peut noter que le GPT possède des outils de jugement dans toutes ses activités, et que le quart des activités du GPP est pourvu d'outils de jugement. Le Tableau 5 permet de voir clairement la relation entre l'intention, la mesure et le jugement en ce qui concerne les outils d'évaluation (outils de jugement). En effet, on peut constater que les guides pédagogiques proposent des activités sans tâches et des tâches sans outils de jugement : mais on les présente tout de même comme des activités d'évaluation formative. De plus, au niveau de l'intention, l'objet d'évaluation n'est pas toujours en lien avec les critères d'évaluation. À ce propos, un des principes retenus demande d'assurer un lien étroit entre ce que vise l'évaluation et ce sur quoi portent les résultats (Comité consultatif mixte, 1993). Il est donc évident que l'objet d'évaluation (le quoi) devient capital parce que sa mention dans l'activité influencerait comme nous l'avons dit plus haut, sur la validité des résultats de l'évaluation formative. De fait, les critères d'évaluation devraient découler de l'opérationnalisation de l'objet d'évaluation en ce sens que le critère ou l'ensemble des critères permettent de recueillir auprès de l'évalué des informations empiriques se rapportant à l'objet d'évaluation. Les problèmes de pertinence et de

désuétude rencontrés ici, couplés avec l'absence ou l'insuffisance des outils d'évaluation dans les guides pédagogiques ne permettent pas de dire que les guides respectent les principes d'évaluation formative associés à cette étape du jugement.

Processus – Communication - Rétroaction

Au sujet de la rétroaction, le préambule du GPP et celui du GPI recommandent ou prescrivent l'usage de rétroactions. Dans le GPP, il est mentionné que « le personnel enseignant peut utiliser les résultats de l'évaluation formative pour différencier l'enseignement et offrir aux élèves une rétroaction immédiate » (GPP, 2010, p. xiv). Dans le GPI, il est indiqué que « la rétroaction de l'enseignante ou de l'enseignant doit être descriptive et précise » (GPI, 2009, p. VE-43). Aucune donnée ne fait allusion à la rétroaction dans les contenus du GPP et du GPI, ni dans l'intégralité du GPT.

Cependant, selon Cizek (2010), il faudrait donner à l'élève une rétroaction spécifique, opportune et non évaluative. Dans la même lancée, Clark (2010) demande de donner à l'élève une rétroaction immédiate sur la qualité de son travail et les possibilités d'amélioration. Pour l'OCDE (2005), il est question de rétroagir sur les performances de l'élève et adapter l'enseignement pour répondre aux besoins identifiés. Les travaux de tous ces auteurs mettent en évidence un certain nombre de caractéristiques associées à la rétroaction en contexte d'évaluation formative. De plus, l'évaluation formative de haute qualité prescrite par Black et William (1998b) semble trouver sous la plume d'Earl et Katz (2006) une réponse. Selon Earl et Katz, « L'évaluation *au service de* l'apprentissage est de grande qualité quand l'enseignant peut s'en servir en vue de prendre des décisions suffisamment précises sur l'apprentissage des élèves pour pouvoir fournir une rétroaction descriptive, et afin de préparer la prochaine étape de l'apprentissage » (p. 31). Ainsi, toute rétroaction en contexte d'évaluation formative devrait être conforme aux caractéristiques citées dans ce paragraphe pour que l'activité respecte les principes d'évaluation formative relatifs à la rétroaction; ce que les résultats peinent à montrer.

Conditions – Souplesse et dynamisme

On retrouve des extraits se rapportant à la souplesse dans cinq activités du GPP; dans une activité du GPI et dans aucune activité du GPT. En voici un exemple : « Les élèves peuvent aussi réaliser l'activité Sciences en action à la maison au moyen d'un ou deux appareils ménagers dont

ils analyseront la perte d'énergie » (GPP, 2010, p. 84). Ici, il y a une souplesse parce qu'il est demandé à l'enseignant de donner à l'élève l'occasion de faire son évaluation formative à la maison, ce qui n'aurait été le cas en contexte sommatif.

Quant au dynamisme de l'évaluation formative, le préambule du GPP stipule que : « L'évaluation *en tant* qu'apprentissage a lieu grâce à une discussion sur les stratégies et les outils d'évaluation appropriés et par la pratique de l'autoévaluation, de l'évaluation par les pairs, de la réflexion et de la métacognition » (GPP, 2010, p. xiv). Dans cet extrait, si l'on s'arrête à la discussion dans laquelle on veut engager les différentes parties prenantes, on a affaire au dynamisme. De fait, une discussion favorise la plupart du temps la compréhension et la réflexion, puis rend plus féconds les apprentissages. Dans le contenu du GPP on retrouve des extraits se rapportant au dynamisme dans cinq activités. Quant au GPI, on retrouve des extraits se portant sur le dynamisme dans trois activités. En voici un exemple : « Si les élèves font l'activité individuellement, discuter des réponses avec la classe » (GPI, 2009, p. 59). Ici, il y a dynamisme dans l'interaction entre l'élève, l'enseignant et toute la classe. Il n'y a pas d'extraits se rapportant au dynamisme dans le GPT.

Conditions – Droit à l'erreur

Le droit à l'erreur est perceptible dans le GPP, dans deux extraits dont celui ci-après :

« Répétez-leur que le processus de conception n'est pas une démarche linéaire, mais un retour constant sur les étapes précédentes après évaluation de l'étape en cours. S'ils se butent à un obstacle à ce stade, les élèves doivent revenir à l'étape de la recherche ou au plan de conception pour procéder aux changements nécessaires. Rassurez-les en leur affirmant que cette démarche est normale. Leur évaluation finale n'en souffrira pas vraiment. Insistez encore une fois sur l'importance de noter toutes les modifications effectuées à toutes les étapes d'exécution » (GPP, 2010, p. 117).

Dans cet extrait le GPP offre à l'élève la possibilité de faire des erreurs sans être pénalisé. Autrement dit, il semble permettre à l'élève d'apprendre de son erreur et revenir à l'étape précédente pour voir s'il réussit à surmonter les obstacles rencontrés. Il n'y a pas d'extraits se rapportant au droit à l'erreur dans le GPI et le GPT.

La souplesse, le dynamisme et le droit à l'erreur sont des concepts qui rendent une activité d'évaluation facile et simple d'utilisation comme le diraient Scallon (1988a) et Trumpower et Sarwar (2010). Ceci est capital selon ces auteurs pour la survie de l'évaluation formative. Par

exemple, la présence de souplesse dans les scénarios d'évaluation formative diminuerait leur rigidité contrairement aux scénarios d'évaluation sommative suivant l'idée de Scallon parce que le processus formatif selon l'OCDE (2005) sécurise les élèves afin que ceux-ci osent prendre des risques et faire des erreurs constructives. Le manque de souplesse, associé au fait que la fonction d'évaluation n'est pas mentionnée au début de l'activité d'évaluation formative, fait en sorte que le scénario d'évaluation prend des allures sommatives, ce qui influence les apprentissages de l'élève. En d'autres termes, les activités analysées respecteraient les principes d'évaluation formative si l'on trouvait dans le Tableau 6, portant sur la quantification des données relatives aux conditions, des valeurs proches de l'effectif n pour la souplesse, le dynamisme et le droit à l'erreur. Ces concepts sont en fait des conditions pour la survie de chaque activité d'évaluation formative face à la pression engendrée par l'évaluation sommative. De plus, si les enseignants se plaignent du manque de temps pour effectuer des tâches d'évaluation formative, la moindre lourdeur dans l'exécution de celles-ci serait décourageante.

Tableau 6 : Quantification des données par dimension – Conditions

Conditions	GPP $n = 40$	GPI $n = 21$	GPT $n = 2$
Souplesse	5	0	0
Dynamisme	5	4	0
Droit à l'erreur	1	0	0
Différenciation	9	0	0
Responsabilisation	7	0	1
Interactivité	0	0	1

n représente le nombre d'unités d'analyse de chaque guide pédagogique

Conditions – Différenciation

On retrouve des extraits relatifs à la différenciation dans neuf activités du GPP; par exemple :

« La répétition des présentations en petites équipes peut mettre les élèves plus à l'aise. Laissez les élèves en FLS travailler avec des camarades se débrouillant mieux en français. Assurez-vous qu'ils participent à la présentation conformément à leurs aptitudes. Une ou un élève dont la compétence langagière est limitée peut communiquer au moyen d'illustrations et de mots simples » (GPP, 2010, p. 41).

« Des élèves éprouveront peut-être de la nervosité à l'idée de présenter leur dispositif devant toute la classe. Permettez-leur d'opter pour un autre mode de présentation ou de remplir une autre tâche (technicienne ou technicien, par exemple) s'ils travaillent par deux ou en équipe. Sinon, ils pourraient effectuer leur présentation devant vous ou un petit groupe en dehors des heures de classe » (GPP, 2010, p. 115);

Ces extraits portent la plupart du temps sur l'élève dont le français est langue seconde, qui aurait des difficultés de compréhension à cause de sa langue maternelle. En ce qui concerne le GPI, son préambule recommande que l'enseignant différencie son évaluation formative par l'usage d'une part, de rétroactions spécifiques pour les élèves ayant des difficultés, et d'autre part, des stratégies variées telles que les productions, les conversations et les observations. Cependant, on ne trouve pas de telles stratégies appliquées spécifiquement aux activités proposées. Il n'y a pas d'extraits portant sur la différenciation dans le GPT.

Notons que le GPP tient compte, dans la plupart des extraits, des élèves dont les capacités linguistiques pourraient les empêcher d'apprendre au même rythme que les autres; ce qui contribue à la promotion de l'égalité de chances pour tous les élèves (Laveault, 2010). Cependant, la différenciation ne s'arrête pas aux élèves dont le français est langue seconde. Que dire de ceux qui sont en difficulté d'apprentissage? Ces derniers méritent tout autant d'être accommodés pendant les activités d'évaluation formative, comme le montre le deuxième extrait ci-dessus. Selon le Comité consultatif mixte (1993), on devrait offrir à chaque élève suffisamment d'occasions de manifester les acquis faisant partie de l'évaluation, laquelle évaluation devrait au sens de Cizek (2010), être fréquente. Tout compte fait, on pourrait s'accorder sur le fait que la disparité et la récurrence des activités d'évaluation formative participent de la différenciation de l'évaluation formative, en ce sens qu'elles aideraient à ce que le plus grand nombre d'élèves possible puissent réussir. Cependant, les tâches d'évaluation ne sont pas si diversifiées que cela dans le GPP. Et dans le GPI, les stratégies d'évaluation formative variées énoncées dans le préambule ne se retrouvent pas dans le contenu. Le GPI et le GPT semblent ne pas différencier les activités d'évaluation formative. Ce qui fait que le concept de différenciation semble partiellement pris en compte dans les guides pédagogiques.

Conditions – Responsabilisation

On retrouve des extraits relatifs à la responsabilisation dans le préambule du GPP et du GPI. Par exemple, on peut lire dans le préambule du GPP :

« La rubrique *Fais un résumé (...)* offre une occasion de révision et d'autoévaluation » (GPP, 2010, p. ix);
« *Perspectives* fournit aux élèves bon nombre d'occasions d'autoévaluation, d'évaluation par leurs pairs et de réflexion sur leur apprentissage. Les rubriques *Vérifie ta compréhension*, *Révision du chapitre* et *Révision de l'unité* contiennent plusieurs questions qui demandent aux élèves de réfléchir en profondeur à leur apprentissage, à la façon dont ils l'ont acquis et aux moyens de l'améliorer » (GPP, 2010, p. xvi).

Quand on lit les extraits ci-dessus, on s'attend à voir dans le contenu du GPP un certain nombre d'éléments; en effet, le GPP propose sept activités autoévaluées assorties d'échelles d'appréciation qui aident l'élève à se responsabiliser dans ses apprentissages et réguler ces derniers par la pratique de l'autoévaluation. Dans le contenu du GPP, on trouve un seul extrait relatif à la métacognition : « Affichez le tableau SVA [ce que je *sais*, ce que je *veux* savoir, ce que j'ai *appris*] élaboré à l'activité *Point de départ* au début de l'unité. Demandez aux élèves de lire les réponses et de réfléchir à ce qu'ils ont appris depuis ce temps » (GPP, 2010, p. 119). Ici en effet, on invite l'élève à réfléchir sur ses apprentissages. Pour le contenu du GPI, il n'y a pas d'extraits relatifs à la responsabilisation. En ce qui concerne le GPT, il est mentionné dans le préambule que « L'élève devra montrer son engagement dans ses apprentissages en : Dressant un portfolio des tâches; (...) Participant aux discussions et à la foire des sciences » (GPT, 2001, p. 19). Ceci est un exemple de données portant sur la responsabilisation de l'élève dont le GPT semble prendre en compte. Ici, la responsabilisation est aussi soulignée par la présence d'une échelle d'appréciation pour l'autoévaluation qui est assortie à une des activités proposées.

Conditions – Interactivité

Les extraits relatifs à l'interactivité ne se retrouvent dans aucune activité du GPP et GPI. En ce qui concerne le GPT, la présence d'une échelle d'appréciation pour l'évaluation par les pairs témoigne de l'interactivité entre les élèves au cours de l'évaluation formative. Cependant, dans la tâche on ne fait aucune allusion à l'évaluation par les pairs.

La responsabilisation et l'interactivité semblent partiellement prises en compte dans les activités proposées dans les guides pédagogiques. Selon Cizek (2010), on devrait fournir à l'élève des autoévaluations et des évaluations par les pairs. Le GPP est pourvu de tâches d'autoévaluation et le GPI en est dépourvu. Il semble en effet surprenant qu'aucune activité autoévaluée ne figure dans un guide pédagogique comme le GPI; et aussi qu'aucune activité d'évaluation par les pairs ne figure dans l'ensemble des guides. Il est vrai qu'on ne s'attend pas à ce que chaque activité comporte un volet d'autoévaluation ou d'évaluation par les pairs, mais on devrait tout de même à

notre sens trouver un certain nombre d'activités qui font appel aux pairs ou à l'évalué comme responsables d'évaluation. Par ailleurs, la présence d'activités d'autoévaluation ne suffit pas pour que le concept de responsabilisation respecte les principes d'évaluation formative qui y sont rattachés; car l'autonomie de l'élève implique aussi une réflexion sur ses apprentissages (Cizek, 2010; Earl et Katz, 2006 et OCDE, 2005). Il s'agit donc de satisfaire aux principes relatifs à la métacognition, ce qui n'est pas toujours le cas dans les guides pédagogiques.

En règle générale, la nonoccurrence des données dans le contenu des activités, reflété dans le Tableau 5 et le Tableau 6, ne signifie pas toujours que les principes relevant de ces données ne sont pas respectés. En fait, certains principes présents dans le préambule du guide pédagogique peuvent tacitement s'appliquer à l'ensemble des activités. C'est pourquoi, par exemple, on ne voit aucune mention de la rétroaction dans les contenus, sans pour autant que leur prise en compte dans les unités d'analyse soit insuffisante ou nulle du fait de leur nonoccurrence dans les activités. Notons aussi que la définition d'un principe que nous avons faite dans le cadre conceptuel reflète ce que nous considérons comme tel, dans cette étude. En d'autres termes, il s'agirait de propositions, recommandations, suggestions, etc. faites aux enseignants. Dans cet ordre d'idées, l'occurrence des principes aurait pu être renforcée de la façon suivante : dans le préambule du GPI il est mentionné que « Les critères d'évaluation **aident** les élèves à concevoir ce qu'est une échelle de rendement et constituent une étape essentielle du processus d'évaluation » (p. VE-43). Cet extrait est une déclaration qui n'invite pas explicitement l'enseignant à faire quoi que ce soit. Par contre, si cette déclaration était transformée en une recommandation telle que « Les élèves **devraient** aussi avoir en leur possession les grilles d'évaluation ou les guides de notation conçus à l'aide de la grille d'évaluation du rendement » (p. VE-43) – les mots clés sont volontairement mis en gras –; là on aurait eu affaire à un principe invitant explicitement l'enseignant à remettre à l'élève les grilles d'évaluation et les guides de notation, par souci dirions-nous de transparence comme le soulignent Forgette-Giroux et Simon (1998) et comme le recommande Gronlund (2003). L'usage de proposition au sens de Legendre (2005) comme « les élèves devraient... » pourrait ainsi accroître l'occurrence des principes et permettre à l'enseignant d'augmenter en qualité l'efficacité des activités d'évaluation formative pour l'amélioration des apprentissages de l'élève.

Quand on regarde les résultats détaillés qui sont présentés dans le Tableau 7, encore aussi présentés dans l'Annexe 3, l'Annexe 4 et l'Annexe 5, on pourrait être en droit dire que le GPI, à titre d'exemple bien entendu, n'aide pas l'enseignant à faire de l'évaluation formative. Nous

soutenons comme Bennett (2011), Black et Wiliam (1998b) et McMillan (2004) que des activités d'évaluation formative de haute qualité devraient respecter ces principes dans leur intégralité. Ce niveau de qualité serait selon McMillan (2004) tributaire de la validité et la fidélité. Bien que ces deux concepts soient associés à tout type d'évaluation, ils peuvent tout de même appuyer la qualité de la prise en compte des principes d'évaluation formative. À ce propos, on peut souligner que les attributs de qualité comme les insuffisances des instruments de cueillette d'information ou la désuétude des ressources, de même que les problèmes d'incohérence entre les objets d'évaluation et les critères d'évaluation affectent la validité des résultats de mesure.

Tableau 7 : Occurrence des données – exemple du GPI

Section du GPP	Activités/rubriques d'évaluation formative suggérée	Processus							Conditions					
		Fonction d'évaluation Responsable évaluation	Objet d'évaluation	Tâche	Instruments	Outils de jugement	Critères d'évaluation	Rétroaction	Souplesse	Dynamisme	Droit à l'erreur	Différentiation	Responsabilisation	Interactivité
	Vue d'ensemble					x	x	x	x				x	
4.1	1. B4 : Point de départ			x										
	2. B6 : Fais le point !			x										
	3. Vérifie et réfléchis													
4.2	4. B10 Point de départ		x							x				
	5. B12 : Fais le point !			x										
	6. Vérifie et réfléchis													
4.3	7. B15 Point de départ			x										
	8. B16 Fais le point !			x										
	9. B17 Fais le point !			x										
	10. B18 Fais le point !			x										
	11. Vérifie et réfléchis													
5.1	12. B24 Point de départ			x										
	13. B26 Fais le point !			x						x				
	14. Vérifie et réfléchis													
5.2	15. B30 Point de départ			x										
	16. B31 Fais le point !			x										
	17. Vérifie et réfléchis													
6.1	18. B36 Point de départ			x						x				
	19. Vérifie et réfléchis													
6.2	20. B41 Point de départ			x						x				
	21. Vérifie et réfléchis													
	Total (hors Vue d'ensemble)	0	0	1	13	0	0	0	0	0	4	0	0	0

x = Occurrence; □ = non-occurrence.

Un autre volet de notre discussion se rapporte au domaine d'étude ciblé avant la recension des écrits. En fait, le domaine circonscrit dans la problématique est l'éducation technologique. Le cadre méthodologique a permis de circonscrire les guides pédagogiques disponibles pouvant être des objets d'étude. En plus, le sujet d'étude choisi à savoir « Les systèmes en action » correspond à la partie du curriculum de sciences et technologie de l'Ontario portant sur la technologie traitant des concepts technologiques. En examinant les données extraites des guides lors de la collecte de données, en ce qui concerne le GPP et le GPI, on peut constater que celles-ci ne touchent pas vraiment la technologie telle que définie par le MÉO (2009), c'est-à-dire basée sur le travail pratique et la réalisation des projets technologiques en vue de produire des biens ou offrir des services. Seul le GPT propose des tâches en lien avec la technologie. Ce constat pourrait interpeller soit les concepteurs des guides pédagogiques, soit le MÉO à intégrer les principes technologiques dans les ouvrages proposés.

2.6. Limites et conclusion

L'une des limites de cette étude porte sur le corpus d'analyse qui pourrait sembler insuffisant du fait du nombre réduit de ressources analysées de surcroît partiellement. Mais du fait que le domaine d'étude a été synthétisé dans la problématique et circonscrit dans la méthodologie, cette limite n'a pas de conséquence sur les résultats de l'étude. Par ailleurs, lors de la validation des instruments de collecte de données, une triangulation des chercheurs aurait permis de mieux affiner les questions guides. Néanmoins, cette étude pilote a donné lieu à un accord interjuge acceptable. Enfin, certaines ressources indexées aux guides pédagogiques analysés n'ont pas été analysées parce que ne faisant pas partie du corpus d'analyse. Peut-être que leur prise en compte aurait permis de trouver des résultats différents. Cependant, cette limite n'affecte pas nos résultats, étant donné la précision de notre question spécifique de recherche qui a orienté cette étude.

Cette étude a permis de partir des principes d'évaluation formative découlant de la recension des écrits pour arriver à une liste de principes d'évaluation formative que nous proposons sous forme de recommandations. Ainsi, pour une prise en compte des principes d'évaluation formative, un guide pédagogique devrait respecter les principes suivants :

1. Préciser les intentions d'évaluation, en l'occurrence la fonction d'évaluation formative, le responsable d'évaluation formative et l'objet d'évaluation formative;

2. Pourvoir les activités d'évaluation formative de tâches et d'instruments spécifiques et variés, qui permettent de mesurer une gamme complète d'habiletés, tout en mettant un accent sur les habiletés supérieures à la mémorisation;
3. Proposer des outils d'évaluation comportant des critères d'évaluation qui soient en lien avec l'objet d'évaluation;
4. Suggérer une rétroaction qui respecte les caractéristiques ci-après, de façon à permettre une régulation de l'enseignement et des apprentissages en cours : descriptive, spécifique, opportune, immédiate, non évaluative et adaptée aux besoins identifiés;
5. Assouplir les scénarios d'évaluation formative et favoriser l'exercice du droit à l'erreur, tout en dynamisant les conditions de mise en œuvre de l'activité;
6. Différencier les activités d'évaluation formative pour inclure l'élève en différentes situations d'apprentissage : difficulté d'apprentissage, minorité linguistique, dispositions psychologiques (par exemple, la nervosité), etc.
7. Responsabiliser l'élève en l'amenant à réfléchir en profondeur sur ses apprentissages, puis favoriser la pratique de l'autoévaluation et de l'évaluation par les pairs.

En guise de conclusion, rappelons que cette étude a pour but d'explorer les guides pédagogiques de sciences et technologie de l'Ontario pour voir dans quelles mesures ils prennent en compte les principes d'évaluation formative dans les activités d'évaluation formative proposées aux enseignants. Malgré la popularité dont jouit l'évaluation formative, comment se fait-il encore aujourd'hui que certains professionnels hésitent à en faire un usage productif dans le sens d'améliorer les apprentissages des élèves? Se pourrait-il que ce problème résulte de manquements quelconques? Le cas échéant, lesquels? La réponse à toutes ces questions est sans doute liée à l'opportunité qui est donnée aux enseignants de mettre en œuvre l'évaluation formative, notamment par l'usage d'un guide pédagogique dont les activités proposées à des fins formatives respectent les principes d'évaluation formative. C'est ce que nous avons essayé de comprendre en effectuant cette étude, dont les résultats tendent à montrer que les principes d'évaluation formative ne sont pas toujours pris en compte dans les guides pédagogiques. Il n'est pas facile de faire ressortir des tendances générales de ces résultats du fait qu'aucun des concepts associés aux principes d'évaluation formative n'a satisfait auxdits principes de façon intégrale. De plus, ces résultats varient beaucoup d'un guide à l'autre.

Les discussions autour des résultats semblent appuyer la tendance vers le non-respect des principes d'évaluation formative dans la plupart des activités d'évaluation à des fins formatives proposées dans les guides pédagogiques analysés. Peut-être est-ce l'une des raisons pour lesquelles l'opérationnalisation de l'évaluation formative rencontre encore aujourd'hui des obstacles. Il est vrai que la mise en œuvre de l'évaluation formative n'est pas aisée pour l'enseignant; cependant, le non-respect des principes d'évaluation formative semble donner raison à Allal et Laveault (2009) qui pensent que l'enseignant ne devrait pas se contenter des pratiques mécaniques et superficielles qui n'ont rien à voir avec l'évaluation formative. Cette étude sensibilise à la fois l'enseignant, le concepteur de guides pédagogiques et le ministère à la prise en compte des principes d'évaluation formative dans les ressources mises à la disposition de l'enseignant. Nous soutenons que de telles ressources intégrant l'ensemble des principes élaborés dans cet article contribueraient à rendre à l'évaluation formative ses lettres de noblesse.

Chapitre 3 : Conclusion générale

Le but de cette thèse sous forme d'article était d'explorer les guides pédagogiques de sciences et technologie de 8^e année de l'Ontario pour voir dans quelles mesures ils intègrent les principes d'évaluation formative dans les activités proposées à des fins formatives aux enseignants. Le premier chapitre était consacré à présentation générale de l'étude. Nous y avons problématisé l'étude, ensuite nous l'avons conceptualisée et enfin opérationnalisée. Dans le deuxième chapitre, nous avons présenté notre article proprement dit, que nous comptons soumettre à une revue arbitrée aux fins d'évaluation et de publication. L'article est complet, mais aussi plus succinct. Dans le présent chapitre, nous présentons les conclusions générales de l'étude.

3.1. Rappel du contexte

Cette étude part du postulat que les guides pédagogiques sont très utilisés par les enseignants (Hasni *et al.*, 2009) et que malgré la mise en valeur des communautés d'Apprentissage professionnel (CAP) sur lesquelles le Ministère de l'Éducation de l'Ontario (MÉO) met un accent particulier, les enseignants ont besoin de ressources pédagogiques qui puissent les aider à effectuer des évaluations formatives de haute qualité (Black et Wiliam, 1998b). Le guide pédagogique a été présenté comme un moyen dont les enseignants peuvent se servir pour mettre en œuvre l'évaluation formative au quotidien. En effet, le rôle du guide pédagogique ne se limite pas à fournir des savoirs disciplinaires, mais aussi des suggestions et des approches pédagogiques pour aider l'enseignant à mieux planifier les activités d'évaluation formative. Dans cette optique, le guide pédagogique a été présenté comme un vecteur d'actualisation des politiques évaluatives et des fruits de la recherche en matière d'évaluation formative (Niclot, 2003). Ceci voudrait dire que les manquements à la pratique de l'évaluation formative peuvent être comblés par un guide pédagogique au fait des avancées scientifiques dans le domaine de l'évaluation formative et pourvu de dispositifs pouvant aider l'enseignant à mettre en œuvre l'évaluation formative.

L'évaluation formative est sous-tendue par des principes qui lui sont propres. Dans la problématique, nous avons mis à l'avant le fait que la sous-représentativité des principes d'évaluation formative est l'un des éléments pouvant faire obstacle à la mise en œuvre de l'évaluation formative dans les pratiques quotidiennes des enseignants (Bennett, 2011 ; Earl et Katz, 2006). Les principes d'évaluation formative issus de la recherche en éducation sont

nombreux. Nous avons trouvé dans les écrits scientifiques recensés de nombreux principes énumérés tant par les auteurs clés du domaine que par d'autres études.

Ceci nous a amené au cadre conceptuel où les concepts émanant de la recension des écrits ont été définis et mis en relation. Ces concepts ont été regroupés sur deux dimensions : le processus d'évaluation et les conditions entourant l'évaluation formative. La définition des concepts et leur mise en relation ont conduit à la question de recherche. Cette question de recherche nous amène à mesurer la prise en compte des principes d'évaluation formative dans les guides pédagogiques afin de comprendre comment ceux-ci aident les enseignants à opérationnaliser cette pratique évaluative.

Pour effectuer cette mesure, deux instruments de collecte de données ont été conçus dans le cadre méthodologique, sur la base des dimensions et des concepts à l'étude, définis et élaborés dans le cadre conceptuel. Ces instruments ont subi plusieurs épreuves de validation avant être utilisés. Il a été question de conformer les instruments aux règles d'élaboration des listes de vérification, puis un exercice d'appariement a permis de valider leur contenu et enfin, une étude pilote s'est avérée utile pour garantir l'exclusivité mutuelle des codes et établir un indice de fidélité interjuge acceptable. C'est également dans le cadre méthodologique que les catégories analytiques et le système de codage ont été énoncés. De fait, les catégories analytiques et les codes étaient arrimés aux dimensions et aux concepts respectivement. Les unités d'étude ont été affinées au regard de la problématique, ce qui a donné lieu à la prise en considération de trois guides pédagogiques : le GPP, le GPI et le GPT. Ces unités d'étude comportaient chacune des sources de données, en l'occurrence un préambule et un contenu. Notons que les données qui ressortent de l'une ou l'autre de ces sources de données n'étaient pas interprétées de la même manière ; par exemple, les données issues du préambule semblaient s'appliquer à l'ensemble de l'unité d'étude tandis que les données extirpées du contenu s'appliquaient le plus souvent à chaque unité d'analyse. Était considérée comme unité d'analyse, toute activité d'évaluation proposée dans le préambule du guide pédagogique à des fins formatives. Les données issues de la collecte de données ont été analysées à l'aide d'une grille d'analyse, laquelle associait les informations contenues dans les unités d'analyse et les données contenues dans les instruments de collecte de données.

Les résultats de cette étude ont permis de se prononcer tantôt en faveur de la prise en compte des principes d'évaluation formative, tantôt contre. La sous-représentativité des principes d'évaluation formative dans le guide pédagogique pourrait expliquer pourquoi certains enseignants éprouvent des difficultés à mettre en pratique l'évaluation formative, ou à ne pas en comprendre le bienfondé sur les apprentissages de leurs élèves. Nos résultats peuvent sembler alarmants parce qu'ils nous amènent à nous questionner sur l'opérationnalisation de l'évaluation formative. Cependant, il ne faut pas perdre de vue que le guide pédagogique est présenté comme un médium parmi d'autres, bien qu'il nous semble être parmi les plus efficaces pour aider les enseignants à effectuer des évaluations formatives de haute qualité.

3.2. Retombées de l'étude

Plusieurs implications d'ordre théoriques et pratiques découlant de cette étude pourront d'une part aider à réfléchir plus en profondeur sur les problèmes soulevés et d'autre part inspirer les émetteurs et les destinataires des guides pédagogiques.

3.2.1. Implications théoriques

Comme mentionné plus haut, au moment de la recherche bibliographique, il nous a été impossible de trouver des articles scientifiques portant sur le domaine du guide pédagogique de sciences en matière d'évaluation formative. En effet, la recension a permis de se rendre compte que loin d'avoir suscité l'intérêt de la recherche, le guide pédagogique semble ne pas avoir fait l'objet de nombreuses études comme le confirment Margolinas et Wozniak (2009), surtout dans le domaine de l'évaluation formative en Ontario. Pourtant les résultats de cette étude montrent que ce domaine de la recherche mérite d'être exploré plus en profondeur, en l'occurrence sur plusieurs axes : un premier axe pourrait consister à répéter l'étude dans des matières autres que les sciences et technologie ; un autre axe pourrait consister à explorer les différents paliers éducatifs tels que les cycles primaire et secondaire. De telles études pourraient aider à voir comment la matière et le niveau éducatif enseignés influencent la conception des guides pédagogiques en matière d'évaluation formative. Donc, cette étude a exposé des lacunes qui pourraient ouvrir des pistes de recherches afin d'enrichir ce domaine.

Pendant cette étude, nous avons donc développé un cadre de référence qui pourrait être étudié plus en profondeur par le milieu académique pour la prise en compte des principes d'évaluation formative non seulement des guides pédagogiques, mais dans tout autre document à

intention formative. De telles études pourraient par ailleurs renforcer la validité des instruments de collecte de données proposés, ou en améliorer la qualité.

Au cours de la collecte de données, nous sommes resté ouvert à l'émergence de nouveaux codes. Cependant, il n'y a rien eu de plus que les codes de départ. Ceci amène éventuellement à se demander si une telle précaution était nécessaire. Qu'aurait vu un autre chercheur, par exemple ? Ainsi, un certain intérêt de recherche pourrait découler de cette situation.

3.2.2. Implications pratiques

Outre le cadre de référence proposé, cette étude permet de constater que les guides pédagogiques de sciences du palier intermédiaire approuvés par le MÉO sont perfectibles en matière de prise en compte des principes d'évaluation formative dans les activités proposées, d'une part. En ce sens, notre étude pourrait être une source d'inspiration pour les concepteurs de guides pédagogiques qui pourraient alors offrir aux enseignants une meilleure guidance dans la mise en œuvre d'activités d'évaluation formative de haute qualité (Black et Wiliam, 1998b). À titre d'exemple, notre étude pourrait offrir aux concepteurs de guides pédagogiques et aux enseignants, des outils de contrôle qui leur permettraient de s'assurer que les activités d'évaluation formative qu'ils ont respectivement créées et modifiées respectent les principes d'évaluation formative. D'autre part, les guides seraient également perfectibles en ce qui concerne l'aspect technologique de la matière tel que défini par le MÉO. Si le MÉO considère la technologie au sens du travail pratique et de la réalisation des projets technologiques, ces dispositions devraient être reflétées dans les guides pédagogiques en vigueur.

Enfin, les instruments que nous proposons aux enseignants, qui dans un esprit socioconstructiviste souhaiteraient concevoir eux-mêmes leurs propres activités d'évaluation formative, pourraient aider ces enseignants à mesurer la prise en compte des principes d'évaluation formative dans lesdites activités. Ce qui leur permettrait évidemment d'aider l'élève à progresser dans ses apprentissages.

3.3. Limites

L'une des limites de cette étude pourrait porter sur le corpus d'analyse. En fait, nous avons analysé trois guides pédagogiques et dans chaque guide pédagogique, environ une unité sur quatre, ce qui pourrait sembler insuffisant. Mais du fait que le domaine d'étude a été synthétisé dans la

problématique et circonscrit dans la méthodologie, cette limite n'a pas de conséquence sur les résultats de l'étude.

Par ailleurs, les instruments de collecte de données conçus pour les besoins de l'étude ont été validés en plusieurs étapes : validité de contenu par exercice d'appariement et pilote par deux chercheurs. Cependant, nous pensons qu'une triangulation du chercheur lors de l'étude pilote aurait permis de mieux affiner les questions guides de la collecte de données. Néanmoins, cette étude pilote a tout de même donné lieu à un accord interjuge acceptable.

Enfin, nous avons constaté que certaines ressources indexées aux guides pédagogiques analysés comme les rubriques se trouvant dans le manuel de l'élève n'ont pas été analysées parce qu'elles ne font pas partie du corpus d'analyse. Peut-être que leur prise en compte aurait permis de trouver des résultats différents. Cependant, cette limite n'affecte pas nos résultats, étant donné la précision de notre question spécifique de recherche.

Références

- Agiro, C. P. (2012). Comparative critical discourse analysis of student and teacher editions of secondary Christian American literature textbooks. *Journal of Research on Christian Education*, 21(3), 211-234.
- Allal, L. (1989). Stratégies d'évaluation formative : conceptions psycho-pédagogiques et modalités d'application. Dans Linda Allal, Jean Cardinet et Philippe Perrenoud (dir). *L'évaluation formative dans un enseignement différencié : actes du colloque à l'Université de Genève, mars 1978* (p.153-183). Berne, Suisse : Peter Lang.
- Allal, L. (1993) Régulations métacognitives : quelle place pour l'élève dans l'évaluation formative. Dans Linda Allal, Daniel Bain, et Philippe Perrenoud (dir.) *Évaluation formative et didactique du français*, (81-98). Neuchâtel et Paris : Delachaux et Niestlé.
- Allal, L. et Laveault, D. (2009). Assessment for Learning, Évaluation-soutien d'Apprentissage. *Mesure et évaluation en éducation*, 32(2), 99-106.
- Allal, L. et Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage : revue de publication en langue française. Dans OCDE. *L'évaluation formative – pour un meilleur apprentissage dans les classes secondaires*. OECD Publishing. Récupéré le 12 novembre 2012 de http://www.oecd-ilibrary.org/education/l-evaluation-formative_9789264007420-fr
- American College Testing. (2008). The Forgotten Middle: Ensuring that all Students Are on For College and Career Readiness before high school. Récupéré le 7 octobre 2013 de <http://www.act.org/research/policymakers/pdf/ForgottenMiddleSummary.pdf>
- Bélaïr, L. (1989). Évaluer l'évaluation formative pour une approche communicative efficace. *Journal de l'immersion*, 13(1), 25-31.
- Bennett, R. E. (2011). Formative assessment: a critical review, *Assessment in Education: Principles, Policy & Practice*, 18(1) 5 – 25.
- Beyer, C. J., Delgado, C., Davis, E. A. et Krajcik, J. (2009). Investigating teacher learning supports in high school biology curricular programs to inform the design of educative curriculum materials. *Journal of Research in Science Teaching*, 46(9), 977–998.
- Black, P., et Wiliam, D. (1998a). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 139-148.

- Black, P. et Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in Education*, 5, 7–74.
- Black, P., Harrison C., Lee, C., Marshall, B., et Wiliam, D. (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86(1), 9-21.
- Brookhart, S. M. (2010). *La rétroaction efficace. Des stratégies pour soutenir les élèves dans leur apprentissage*. Québec, QC : Chenelière.
- Bucheton, D. (1999). Les manuels : Un lien entre l'école, la famille, l'élève et les savoirs. Dans S. Plane (dir.) *Manuels et enseignement du français* (p. 41-50). Caen, France : CRDP Basse-Normandie.
- Buck, G. et Trauth-Nare, A. E. (2009). Preparing teachers to make the formative assessment process integral to science teaching and learning. *Journal of Science Teacher Education*. 20(5), 475–494.
- Cizek, G. J. (2010). An introduction to formative assessment. History, characteristics, and challenges. Dans H. L. Andrade et G. J. Cizek (dir.), *Handbook of formative assessment*. (p. 3-17). New York, NY: Routledge.
- Clark, I. (2010). Formative assessment: There is nothing so practical as a good theory. *Australian Journal of Education*. 54(3), 341-352.
- Comité consultatif mixte. (1993). *Principes d'équité relatifs aux pratiques d'évaluation des apprentissages scolaires au Canada*, Edmonton, AB. Récupéré le 16 novembre 2012 de http://www2.education.ualberta.ca/educ/psych/crame/files/fr_princ.pdf
- Conseil des ministres de l'éducation du Canada (2004). Quality Education for All Young People: Challenges, Trends, and Priorities, Report of Canada. Récupéré le 26 septembre 2013 du site http://www.cmec.ca/Publications/Lists/Publications/Attachments/66/47_ICE_report2.en.pdf
- Davis, E. A. (2006). Preservice elementary teachers' critique of instructional materials for science. *Science & Education*, 90(2), 348–375.
- De Ketele, J.-M. et Gérard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et évaluation en éducation*, 28(3), 1-26.

- Doyle, W. (1990). Classroom knowledge as a foundation for teaching. Dans S. Tozer, T. H. Anderson & B. B. Armbruster (Éds.), *Foundational studies in teacher education: Reexamination* (p. 49–60). New York, NY : Teachers College Press.
- Earl, L. et Katz, S. (2006). *Repenser l'évaluation en classe en fonction des buts visés* (2^e Éd.) Protocole de l'Ouest et du Nord canadiens de collaboration concernant l'éducation. Récupéré le 15 novembre 2012 du site http://www.edu.gov.mb.ca/m12/frpub/me/docs/repenser_eval/
- Feldman, A., & Capobianco, B. M. (2008). Teacher learning of technology enhanced formative assessment. *Journal of Science Education and Technology*, 17, 82–99.
- Forgette-Giroux, R, et Simon, M. (1998). L'application du dossier d'apprentissage à l'université. *Mesure et évaluation en éducation*, 20(3), 85-103.
- Gohier, C. (2011). Le cadre théorique. Dans T. Karsenti et L. Savoie-Zajc (dir.), *La recherche en éducation. Étapes et approches* (p. 83 – 108). Saint-Laurent, QC : ERPI.
- Gronlund, E. N. (2003). *Assessment of Student Achievement*. Boston, MA : Allyn & Bacon.
- Halverson, R., Grigg, J., Prichett, R. et Thomas, C. (2007). The new instructional leadership: Creating data-driven instructional systems in school. *Journal of School Leadership*, 17(2), 159-194.
- Hasni, A., Moresoli, C., Samson, G. et Owen, M.-È. (2009). Points de vue d'enseignants de sciences au premier cycle du secondaire sur les manuels scolaires dans le contexte de l'implantation des nouveaux programmes au Québec. *Revue des sciences de l'éducation*, 35(2), 83-105.
- Hattie, J. et Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112.
- Herman, J. L., Osmundson, E. et Silver, D. (2010). *Capturing quality in formative assessment practice: measurement challenges. CRESST Report 770*. Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Récupéré le 12 septembre 2013 de <http://files.eric.ed.gov/fulltext/ED512648.pdf>

- House, A.E., House, B.J. et Campbell, M.B. (1981). Measures of interobserver agreement: Calculation formulas and distribution effects. *Journal of Behavioral Assessment*, 3(1), 37-57.
- International Technology Education Association. (2000). Standards for Technological Literacy : Content for the Study of Technology. Reston, VA : auteur.
- Jalbert, P., et Munn, J. (2001). L'évaluation d'hier et celle de demain. *Vie Pédagogique*, 120, 48-51.
- Kesidou, S., et Roseman, J. E. (2002). How well do middle school science programs measure up? Findings from project 2061's curriculum review. *Journal of Research in Science Teaching*, 39(6), 522-549.
- L'Écuyer, R. (1987). L'analyse de contenu : notion et étapes. Dans J.-P. Deslauriers, (dir.), *Les Méthodes de la recherche qualitative* (p. 49-65). Ste-Foy, QC : Presses de l'Université du Québec.
- Laurier, M., Tousignant, R., et Morissette, D. (2005). *Les principes de la mesure et de l'évaluation des apprentissages*. Montréal, QC : Gaëtan Morin.
- Laveault, D. (2010). Classroom Assessment in Policy Context (French Sources). Dans P. Peterson, E. Baker, B. McGaw, (dir.), *International Encyclopedia of Education*, vol. 3 (p. 432-437). Oxford, England : Elsevier.
- Legendre, R. (2005). *Dictionnaire actuel de l'éducation*. Montréal, QC : Guérin.
- Lin, S.-F, Chang, W.-H. et Cheng, Y.-J. (2010). The perceived usefulness of teachers' guides for science teachers. *International Journal of science and Mathematics Education*, 9, 1367-1389.
- Lynch, S., Kuipers, J., Pyke, C., et Szesze, M. (2005). Examining the effects of a highly rated science curriculum unit on diverse students: Results from a planning grant. *Journal of Research in Science Teaching*, 42(8), 912-946.
- Margolinas, C. et Wozniak, F. (2009). Usage des manuels dans le travail de l'enseignant : l'enseignement des mathématiques à l'école primaire. *Revue des sciences de l'éducation*, 35(2), 59-82.

- McManus, S. (2008). *Attributes of effective formative assessment*. Washington, DC: Council for Chief State School Officers. Récupéré le 15 avril 2015 du site <http://www.ccsso.org/publications/details.cfm?PublicationID=362>
- McMillan, J. (2004). *Classroom assessment, principles and practice for effective instruction*. Boston, MA : Pearson.
- Ministère de l'Éducation de l'Ontario, (2010). *Faire croître le succès : évaluation et communication du rendement des élèves fréquentant les écoles de l'Ontario*. ON : Imprimeur de la Reine pour l'Ontario.
- Ministère de l'Éducation de l'Ontario. (2009). *Curriculum de l'Ontario 9e et 10e année Éducation technologique*. ON : Imprimeur de la Reine pour l'Ontario.
- Morissette, J. (2009). *Manières de faire l'évaluation formative des apprentissages selon un groupe d'enseignantes du primaire: une perspective interactionniste* (Thèse de doctorat, Université Laval, Laval, QC, Canada). Récupéré du site de l'université : www.theses.ulaval.ca/2009/25992/25992.pdf.
- Mottier-Lopez, L. (2007). Régulations interactives situées dans des dynamiques de microculture de classe. *Mesure et Évaluation en Éducation*, 30(2), 23-47.
- Niclot, D. (2003). Et si les manuels scolaires étaient, par défaut, un outil de professionnalisation des enseignants ? Dans G. Baillat, P. A. Martin et D. Niclot (dir.), *Vers quelle professionnalité enseignante en France et au Québec ?* Paris, France : Centre national de documentation pédagogique (CNDP).
- OCDE (2005). *L'évaluation formative : Pour un meilleur apprentissage dans les classes secondaires*. OECD Publishing. DOI : 10.1787/9789264007420-fr
- Opoku-Amankwa, K., Brew-Hammond, A., et Kofigah, F. E. (2011). What is in a textbook? Investigating the language and literacy learning principles of the 'Gateway to english' textbook series. *Pedagogy, Culture & Society*, 19(2), 291-310. doi: 10.1080/14681366.2011.582264
- Popham, W.J. (2008). *Transformative assessment*. Alexandria, VA: ASCD.
- Richard, S. (2006). L'analyse de contenu pour la recherche en didactique de la littérature. Le traitement de données quantitatives pour une analyse qualitative : parcours d'une approche mixte. *Recherches Qualitatives*. 26(1), 181-207.

- Reys, B. J., Reys, R. E., et Chávez, O. (2004). Why mathematics textbooks matter. *Educational Leadership*, 61, 61–66.
- Savoie-Zajc, L. et Karsenti, T. (2011). La méthodologie. Dans T. Karsenti et L. Savoie-Zajc (dir.), *La recherche en éducation. Étapes et approches* (p. 99 – 122). Saint-Laurent, QC : ERPI.
- Scallon, G. (1988a). *L'évaluation formative des apprentissages. Tome II : l'instrumentation*. Québec, QC : Presses de l'université Laval.
- Scallon, G. (1988b). *L'évaluation formative des apprentissages. Tome I : La réflexion*. Québec, QC : Presses de l'université Laval.
- Scallon, G. (1999). L'harmonisation des fonctions formative et sommative de l'évaluation. Récupéré le 10 décembre 2015 du site : <http://www.fse.ulaval.ca/Gerard.Scallon/fascicules/harmoni.pdf>
- Scallon, G. (2000). *L'évaluation formative*. Saint-Laurent, QC: Renouveau Pédagogique.
- Scallon, G. (2004). L'évaluation des apprentissages dans une approche par compétences. Saint-Laurent, QC : Renouveau pédagogique.
- Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify a theory. *British Educational Research Journal*, 27, 615–631.
- Trumpower, D. L., et Sarwar, G. S. (2010). Formative structural assessment: Using concept maps as assessment for learning, *Concept Maps: Making Learning Meaningful. Actes de la 4e International Conference on Concept Mapping de Viña del Mar, Chile*, 132-136. Récupéré le 30 avril 2014 du site <http://cmc.ihmc.us/cmc2010papers/cmc2010-214.pdf>
- Vygotsky, L. S. et Cole, M. (dir) (1978). *Mind in society : the development of higher psychological processes*. Cambridge, MA : Havard University Press.

Annexes

Annexe 1 : Tableau synthèse des principes d'évaluation formative²

Cizek (2010)	Clark (2010)	Comité consultatif mixte (1993)	McMillan (2004)	OCDE (2005)	Trumppower et Sarwar (2010)
1. Faire en sorte que l'élève prenne en charge son apprentissage	1. Engager les élèves à comprendre clairement ce qu'ils apprennent et ce qu'on attend d'eux.	1. Les méthodes d'évaluation devraient être adaptées aux buts de l'évaluation et à son contexte général.	1. Établir une évaluation de classe de haute qualité en termes de justesse des méthodes d'évaluation, de validité, de fidélité, d'impartialité, etc.	1. Instaurer une culture de classe qui encourage l'interaction et l'utilisation des outils d'évaluation.	1. Le processus d'évaluation formative devrait permettre de mesurer les apprentissages pertinents.
2. Encourager les élèves à veiller sur leurs propres progressions vers les objectifs d'apprentissage	2. Donner aux élèves une rétroaction immédiate sur la qualité de leur travail et sur les possibilités d'amélioration.	2. On devrait offrir à tous les élèves suffisamment d'occasions de manifester les connaissances, les habiletés, les attitudes et les comportements qui font l'objet de l'évaluation.	2. Évaluer le progrès des élèves pendant les unités de cours.	2. Rétroagir sur les performances de l'élève et adapter l'enseignement pour répondre aux besoins identifiés.	2. Identifier les forces et les faiblesses de l'élève.
3. Fournir des évaluations fréquentes, y compris les évaluations par les pairs et les autoévaluations et des évaluations intégrées aux activités d'apprentissage.	3. Prodiger aux élèves des conseils pour soutenir leur progrès.	3. Les rapports d'évaluation devraient être clairs, précis et avoir une valeur pratique pour les personnes à qui ils s'adressent.	3. Effectuer de l'évaluation de performance et du dossier d'apprentissage (portfolio) afin de mesurer les habiletés de compréhension profonde et de la pensée critique.	3. Impliquer activement les élèves dans le processus d'apprentissage.	3. Fournir une rétroaction efficace.
4. Inclure de la rétroaction qui soit spécifique, opportune, axée sur les objectifs d'apprentissage et non évaluative, et qui puisse fournir à l'élève des occasions de réviser et améliorer son travail, puis approfondir sa compréhension.	4. Impliquer les élèves dans les décisions relatives aux étapes subséquentes.				4. Être simple d'utilisation.
5. Promouvoir la métacognition et la réflexion de l'élève sur son propre travail	5. Informer les élèves sur la disponibilité du soutien et leur en procurer en cas de besoin.				
	6. Engage les élèves à construire leurs propres connaissances et à réfléchir sur leurs apprentissages.				
	7. Amener les élèves à être plus responsables de leurs apprentissages et participer au processus d'apprentissage.				

²Les thèmes abordés dans le tableau ont été librement traduits, le cas échéant en français, à partir de leur langue originale.

Annexe 2 : Mise en relation et synthèse des principes

Thème	Principes retenus applicables au guide pédagogique ³	Synthèse des principes
Intention	<ul style="list-style-type: none"> - Évaluer le progrès des élèves pendant les unités de cours. 	Faire des évaluations formatives pendant les unités de cours en précisant la fonction, l'objet, le responsable et le moment.
Mesure	<ul style="list-style-type: none"> - Le processus d'évaluation formative devrait permettre de mesurer les apprentissages pertinents (le quoi). - Effectuer de l'évaluation de performance et du dossier d'apprentissage (portfolio) afin de mesurer les habiletés de compréhension profonde et de la pensée critique (le quoi). - On devrait offrir à tous les élèves suffisamment d'occasions de manifester les connaissances, les habiletés, les attitudes et les comportements qui font l'objet de l'évaluation. 	<p>Faire en sorte que les activités d'évaluation permettent de mesurer les habiletés au-delà de la simple connaissance (progrès, pensée critique).</p> <p>Utiliser une variété d'outils d'évaluation afin de répondre aux styles d'apprentissage diversifiés des élèves.</p> <p>Utilisation d'outils qui permettent de mesurer des apprentissages pertinents (en termes de représentativité du contenu), <i>p. ex.</i> : dossier d'apprentissage, etc.</p>
Jugement	<ul style="list-style-type: none"> - Identifier les forces et les faiblesses de l'élève. - L'évaluation doit être facile à noter et à interpréter [faisabilité]. 	<p>Appliquer les règles d'interprétation propres à l'évaluation formative, soit :</p> <ul style="list-style-type: none"> - analytique - descriptive - qualitative - non comptabilisée
	<ul style="list-style-type: none"> - L'évaluation formative requiert la détermination à priori des critères de réussite. 	Expliciter les critères de réussite de l'évaluation formative.
Communication	<ul style="list-style-type: none"> - Donner aux élèves une rétroaction immédiate sur la qualité de leur travail et sur les possibilités d'amélioration. - Fournir une <u>rétroaction</u> efficace. - L'évaluation formative requiert une rétroaction aux élèves qui mette un accent sur leurs forces et les points à améliorer. - Rétroagir sur les performances de l'élève et adapter l'enseignement pour répondre aux besoins identifiés. 	<p>Orienter l'acte d'évaluation vers la régulation de l'apprentissage à l'aide d'une rétroaction axée sur l'ajustement de :</p> <ul style="list-style-type: none"> - l'enseignement - l'apprentissage.
Souplesse et dynamisme	<ul style="list-style-type: none"> - L'évaluation formative devrait être simple. 	<p>Assurer une certaine souplesse dans l'évaluation formative de manière à adopter au besoin des changements quant :</p> <ul style="list-style-type: none"> - au moment - au lieu - aux moyens (par exemple, usage de documents de référence pendant l'évaluation)
Droit à l'erreur	<ul style="list-style-type: none"> - Instaurer une culture de classe qui encourage l'interaction et l'utilisation des outils d'évaluation. 	Sécuriser les élèves (en ne comptabilisant pas les résultats par exemple) pour que ceux-ci osent prendre des risques et faire des erreurs en classe.
Différenciation	<ul style="list-style-type: none"> - L'évaluation formative requiert l'utilisation d'une variété d'outils d'évaluation. - On devrait offrir à tous les élèves suffisamment d'occasions de manifester les connaissances, les 	Faire en sorte que la tâche d'évaluation puisse permettre de prendre en compte les différences individuelles des élèves.

³ Les termes soulignés ou dans les crochets renvoient aux dimensions auxquelles les principes multidimensionnels font individuellement référence.

Thème	Principes retenus applicables au guide pédagogique ³	Synthèse des principes
	habiletés, les attitudes et les comportements qui font l'objet de l'évaluation.	
Responsabilisation	<ul style="list-style-type: none"> - Engager les élèves à comprendre clairement ce qu'ils apprennent et ce qu'on attend d'eux. - Impliquer les élèves dans les décisions relatives aux étapes subséquentes. - Impliquer activement les élèves dans le processus d'apprentissage. - Engager les élèves à comprendre clairement ce qu'ils apprennent et ce qu'on attend d'eux. (le quoi) 	<p data-bbox="906 302 1333 323">Demander aux élèves de faire des autoévaluations.</p> <p data-bbox="906 359 1398 409">Amener les élèves à être des évaluateurs autonomes en les impliquant :</p> <ul style="list-style-type: none"> - dans les décisions concernant leur progression - dans la réflexion sur leurs apprentissages.
Interactivité	<ul style="list-style-type: none"> - Instaurer une culture de classe qui encourage l'interaction et l'utilisation des outils d'évaluation. - Fournir des évaluations fréquentes, y compris les évaluations par les pairs, les autoévaluations et des évaluations intégrées aux activités d'apprentissage (voir thème Intention). - Fournir des évaluations fréquentes, y compris les évaluations par les pairs, les autoévaluations et des évaluations intégrées aux activités d'apprentissage. 	<p data-bbox="906 604 1414 676">Favoriser les échanges entre les élèves à travers l'évaluation par les pairs en les amenant à utiliser des outils d'évaluation.</p>

Annexe 3 : Occurrence des données par guide pédagogique - GPP

Section du GPP	Activités/rubriques d'évaluation formative suggérée	Processus								Conditions					
		Fonction d'évaluation	Responsable évaluation	Objet d'évaluation	Tâche	Instruments	Outils de jugement	Critères d'évaluation	Rétroaction	Souplesse	Dynamisme	Droit à l'erreur	Différentiation	Responsabilisation	Interactivité
	Préambule				x	x			x		x			x	
1.1	1. Vérifie ta compréhension														
1.2	2. Vérifie ta compréhension														
	3. Sciences en action			x	x	x	x	x							
	4. Info techno														
1.4	5. Sciences en action	x			x						x				
	6. Vérifie ta compréhension														
1.5	7. Prononce-toi sur un enjeu	x	x	x	x	x	x	x			x		x	x	
Chap.1	8. En résumé	x	x	x	x								x		
	9. Révision du chapitre			x											
2.1	10. Activité de fin d'unité	x			x										
	11. Vérifie ta compréhension														
2.2	12. Activité de fin d'unité	x			x										
	13. Sciences en action	x			x			x	x						
	14. Vérifie ta compréhension														
2.3	15. Sciences en action : Calculer ...	x			x						x				
	16. Sciences en action : Les			x	x	x				x					
	17. Vérifie ta compréhension														
2.4	18. Mène une expérience	x	x					x	x						x
2.5	19. Sciences en action : Calculer ...	x			x										
	20. Vérifie ta compréhension														
	21. Géniales, les sciences !														
2.6	22. Mène une expérience	x	x	x				x	x				x	x	
Chap. 2	23. En résumé	x	x	x	x					x			x		
	24. Révision du chapitre			x											
3.1	25. Sciences en action : Détecter ...				x					x					
	26. Vérifie ta compréhension														
	27. Sciences en action : Évaluer ...			x	x	x	x	x							
3.2	28. Sciences en action	x			x			x	x						
	29. Vérifie ta compréhension														
3.3	30. Vérifie ta compréhension														
3.4	31. Mène une expérience	x	x	x				x	x		x		x	x	
3.5	32. Résous un problème	x	x	x				x	x	x		x	x	x	
3.6	33. Sciences en action	x			x										
	34. Vérifie ta compréhension														
3.7	35. Prononce-toi sur un enjeu	x	x	x				x	x				x	x	
	36. Sciences appliquées														
Chap. 3	37. En résumé	x	x	x	x								x		
	38. Révision du chapitre			x											
	39. Activité de fin d'unité	x	x	x	x	x	x	x		x		x	x		
	40. Révision de l'unité		x	x	x	x					x			x	
	Total (hors Préambule)	1	19	17	23	9	11	11	0	5	5	2	9	7	0

☒ = Occurrence ; ☐ = non-occurrence.

Annexe 4 : Occurrence des données par guide pédagogique - GPI

Section du GPP	Activités/rubriques d'évaluation formative suggérée	Processus							Conditions					
		Fonction d'évaluation Responsable évaluation	Objet d'évaluation	Tâche	Instruments	Outils de jugement	Critères d'évaluation	Rétroaction	Souplesse	Dynamisme	Droit à l'erreur	Différentiation	Responsabilisation	Interactivité
	Vue d'ensemble					x	x	x	x				x	
4.1	22. B4 : Point de départ			x										
	23. B6 : Fais le point !			x										
	24. Vérifie et réfléchis													
4.2	25. B10 Point de départ		x							x				
	26. B12 : Fais le point !			x										
	27. Vérifie et réfléchis													
4.3	28. B15 Point de départ			x										
	29. B16 Fais le point !			x										
	30. B17 Fais le point !			x										
	31. B18 Fais le point !			x										
	32. Vérifie et réfléchis													
5.1	33. B24 Point de départ			x										
	34. B26 Fais le point !			x						x				
	35. Vérifie et réfléchis													
5.2	36. B30 Point de départ			x										
	37. B31 Fais le point !			x										
	38. Vérifie et réfléchis													
6.1	39. B36 Point de départ			x						x				
	40. Vérifie et réfléchis													
6.2	41. B41 Point de départ			x						x				
	42. Vérifie et réfléchis													
	Total (hors Vue d'ensemble)	0	0	1	13	0	0	0	0	0	4	0	0	0

x = Occurrence ; □ = non-occurrence.

Annexe 5 : Occurrence des données par guide pédagogique - GPT

Section du GPP	Activités/rubriques d'évaluation formative suggérée	Processus								Conditions				
		Fonction d'évaluation Responsable évaluation	Objet d'évaluation	Tâche	Instruments	Outils de jugement	Critères d'évaluation	Rétroaction	Souplesse	Dynamisme	Droit à l'erreur	Différentiation	Responsabilisation	Interactivité
	Survol												x	
2	1. Plus efficace !	x		x		x	x						x	x
4	2. Fini les obstacles !	x		x		x	x							x
	Total (hors Survol)	0	2	0	2	0	0	2	0	0	0	0	1	2

x = Occurrence ; □ = non-occurrence.

Annexe 6 : Exemples d'extraits tirés des guides pédagogiques analysés

	Code	Exemple	Remarques éventuelles
Fonction de l'évaluation	GPP	« Cette activité diagnostique vous permet d'évaluer les connaissances antérieures des élèves sur les systèmes, les forces et les mécanismes, acquises les années précédentes » (GPP, 2010, p. 18).	
	GPI	Non spécifiée	
	GPT	Non spécifiée	
Responsable de l'évaluation	GPP	« Vous pouvez recueillir le DR 1.5-1 et évaluer les réponses des élèves au moyen de la grille d'évaluation 7, 'Prononce-toi sur un enjeu' » (GPP, 2010, p. 40).	Dans cet extrait, on ne dit pas explicitement qui est le responsable d'évaluation. On peut juste inférer que c'est l'enseignant à qui le guide pédagogique s'adresse
	GPI	Non spécifié	
	GPT	Non spécifié	
Objet d'évaluation	GPP	« Vous pouvez également leur demander de former des équipes pour la préparation d'une présentation. Vous évalueriez la qualité de leurs explications et de leurs idées au moyen de la Grille d'évaluation 1, «Connaissance et compréhension», et de la Grille d'évaluation 3, «Communication» (GPP, 2010, p. 40).	Dans cet exemple, l'objet d'évaluation est l'évaluation la qualité des explications et des idées des élèves.
	GPI	« Cette activité permet d'évaluer la compréhension des élèves au sujet du travail » (GPI, 2009, p. 15)	
	GPT	Non spécifié	
Tâches	GPP	« Plusieurs occasions d'évaluation formative sont intégrées dans <i>Perspectives</i> sous forme de tâches de performance (observer les élèves lors de la démonstration de leurs habiletés et connaissances), d'évaluation de la communication orale (écouter les élèves et discuter avec eux au fil de leur apprentissage) et de travaux écrits (utiliser des stratégies comme les questionnaires demandant une réponse écrite, les jeux-questionnaires et les rapports écrits pour vérifier la compréhension des élèves) » (GPP, 2010, p. xvi).	
	GPI	« B12 Fais le point ! : Calculer le travail <ul style="list-style-type: none"> • Dire aux élèves de relire leurs notes et la section Calculer le travail avant de répondre à ces questions. • Permettre aux élèves de se servir d'une calculatrice pour répondre aux questions. Leur rappeler de commencer par écrire la formule, puis de faire les substitutions à l'aide des unités appropriées. Réponses aux questions 1. 30 J 2. 8,0 J » (GPI, 2009, p. 19).	Dans cet extrait, on peut voir des instructions pour effectuer la tâche d'évaluation formative, mais il n'y a pas de tâche à proprement parler.
	GPT	« Poser des questions pour guider la réflexion de l'équipe. Observer l'habileté de l'élève à mesurer, à noter ses données et à faire ses calculs » (GPT, 2001, p. 19).	
Instruments	GPP	Compris.	(Extrait trop long pour être inséré ici).

Code	Exemple	Remarques éventuelles	
GPI	« L'évaluation équilibrée Pour répondre aux différences existant entre leurs élèves, les enseignants doivent faire appel à diverses stratégies d'évaluation de façon à obtenir des données valables sur le rendement de chacun. Durant la planification de l'évaluation, l'enseignante ou l'enseignant établira donc un équilibre entre les productions, les conversations et les observations » (GPI, 2009, p. VE-45).	Cet extrait suggère aux enseignants l'usage d'instruments variés.	
GPT	Non spécifiés		
Outils de jugement	GPP	« Les <i>Outils d'évaluation</i> contenus dans le présent <i>Guide d'enseignement</i> comprennent les <i>Grilles d'évaluation</i> , les <i>Résumés de l'évaluation</i> (visant le rendement individuel de l'élève) et les <i>Listes de vérification de l'autoévaluation</i> » (GPP, 2010, p. xvii).	Cet extrait nous dit explicitement que le GPP regorge d'outils d'évaluation, et nous précise également la nature des outils d'évaluation que nous pouvons trouver dans cette ressource.
	GPI	« La grille d'évaluation du rendement (...) doit être utilisée de façon équilibrée au moment d'observer ou d'évaluer le travail des élèves. Même si ces compétences sont larges et interdépendantes, elles devraient faire l'objet d'un enseignement explicite et d'une utilisation éclairée » (GPI, 2009, p. VE-44). « Les critères d'évaluation aident les élèves à concevoir ce qu'est une échelle de rendement et constituent une étape essentielle du processus d'évaluation. Les élèves devraient aussi avoir en leur possession les grilles d'évaluation ou les guides de notation conçus à l'aide de la grille d'évaluation du rendement. La participation des élèves à l'élaboration des grilles d'évaluation est particulièrement souhaitable. Elle sera encouragée par l'utilisation des grilles d'évaluation et des listes de contrôle en ligne et sur cédérom correspondant à chaque module de la présente ressource » (GPI, 2009, p. VE-43).	
	GPT	Les outils de jugement sont présents en la grille B et la grille C.	(Extrait trop long pour être inséré ici).
Critères d'évaluation	GPP	Compris	(Extrait trop long pour être inséré ici).
	GPI	« L'étude de copies types d'élèves peut aussi les aider à comprendre les critères d'évaluation... » (GPI, 2009, p. VE43).	
	GPT	Compris	(Extrait trop long pour être inséré ici).
Rétroaction	GPP	« Le personnel enseignant peut utiliser les résultats de l'évaluation formative pour différencier ses pratiques d'enseignement de façon à mieux refléter les besoins des élèves en matière d'apprentissage. Ce type d'enseignement fournit aux élèves une rétroaction immédiate qui favorise la réussite » (GPP, 2010, p. xiv).	
	GPI	« Écrite ou verbale, la rétroaction de l'enseignante ou de l'enseignant doit être descriptive et précise. (...) les enseignants doivent indiquer avec précision aux élèves les éléments qu'ils doivent améliorer et proposer des moyens pour faire progresser ceux d'entre eux qui ont de la difficulté. (...) lorsque les enseignants attribuent des notes sans écrire de remarques ou des notes accompagnées de remarques, les résultats ne sont pas meilleurs. La recherche indique toutefois que le fait de communiquer des remarques sans donner de notes améliore grandement le rendement des élèves (BLACK et WILIAM 1998 ; 2001) » (GPI, 2009, p. VE-43).	L'extrait ci-dessus rappelle à l'enseignant comment faire une rétroaction et quel type de rétroaction serait privilégiée dans le cadre d'une activité d'évaluation formative.

Code	Exemple	Remarques éventuelles
	GPT Non précisée.	
Souplesse	GPP « Les élèves peuvent réaliser l'activité Science en action à la maison » (GPP, 2010, p. 84).	Dans cet extrait, il y a une souplesse quant au lieu où l'activité va être réalisée.
	GPI « La participation des élèves à l'élaboration des grilles d'évaluation est particulièrement souhaitable » (GPI, 2009, p. VE-43).	Participer signifie aussi engendrer d'éventuelles modifications, ce qui concourt à rendre l'activité souple. Nous n'avons pas trouvé des extraits dans le contenu.
	GPT Non précisée	
Dynamisme	GPP « L'évaluation <i>en tant qu'apprentissage</i> a lieu grâce à une discussion sur les stratégies et les outils d'évaluation appropriés et par la pratique de l'autoévaluation, de l'évaluation par les pairs, de la réflexion et de la métacognition » (GPP, 2010, p. xiv).	En effet, la discussion dans laquelle on veut engager les différentes parties prenantes participe du dynamisme de l'évaluation formative.
	GPI En ce qui concerne le dynamisme, nous avons trouvé des extraits s'y rapportant dans trois activités. En voici un exemple : « Si les élèves font l'activité individuellement, discuter des réponses avec la classe » (GPI, 2009, p. 59).	Ici il y a dynamisme dans l'interaction entre les élèves et la classe.
	GPT Non précisée	
Droit à l'erreur	GPP « Répétez-leur que le processus de conception n'est pas une démarche linéaire, mais un retour constant sur les étapes précédentes après évaluation de l'étape en cours. S'ils se butent à un obstacle à ce stade, les élèves doivent revenir à l'étape de la recherche ou au plan de conception pour procéder aux changements nécessaires. Rassurez-les en leur affirmant que cette démarche est normale. Leur évaluation finale n'en souffrira pas vraiment. Insistez encore une fois sur l'importance de noter toutes les modifications effectuées à toutes les étapes d'exécution » (GPP, 2010, p. 117).	Dans cet extrait le GPP offre à l'élève la possibilité de faire des erreurs sans être pénalisé.
	GPI Non précisée	
	GPT Non précisée	
Différenciation	GPP « La répétition des présentations en petites équipes peut mettre les élèves plus à l'aise. Laissez les élèves en FLS travailler avec des camarades se débrouillant mieux en français. Assurez-vous qu'ils participent à la présentation conformément à leurs aptitudes. Une ou un élève dont la compétence langagière est limitée peut communiquer au moyen d'illustrations et de mots simples » (GPP, 2010, p. 41).	Dans cet exemple, on peut se rendre compte qu'il est suggéré d'adapter les moyens d'évaluation au bagage de l'élève en difficulté d'expression due à la langue maternelle de l'élève.
	GPI « les enseignants doivent indiquer avec précision aux élèves les éléments qu'ils doivent améliorer et proposer des moyens pour faire progresser ceux d'entre eux qui ont de la difficulté » (GPI, 2009, p. VE-43)	Ici, on indique quel type de rétroaction serait privilégié dans le cadre d'une différenciation de l'évaluation.
	GPT Non décelée	
Responsabilisation	GPP « La rubrique <i>Fais un résumé (...)</i> offre une occasion de révision et d'autoévaluation » (GPP, 2010, p. ix) ;	
	GPI « L'évaluation des apprentissages aide les élèves à mieux se connaître en s'autoévaluant, en évaluant leurs pairs et en se	Dans cet extrait, il est indiqué qu'une période d'étude et de réflexion devrait aider les élèves

Code	Exemple	Remarques éventuelles
	fixant des objectifs. La période d'étude et de réflexion qui leur est accordée à la fin du cours les aide non seulement à comprendre les derniers concepts étudiés, mais aussi à reconnaître leurs points forts et leurs points faibles ainsi que leurs progrès. » (GPI, 2009, p. VE-46).	dans leurs apprentissages. Par ailleurs, il est mentionné autre part dans le préambule que les copies types d'élèves leur permettent de s'autoévaluer et évaluer leurs pairs. Par contre, dans le contenu du GPI, nous n'avons pas trouvé de mention à l'autoévaluation et à l'évaluation par les pairs.
	GPT Le GPT stipule que « L'élève devra montrer son engagement dans ses apprentissages en : Dressant un portfolio des tâches ; (...). Participant aux discussions et à la foire des sciences » (GPT, 2001, p. 19).	Ceci est un exemple de données portant sur la responsabilisation de l'élève dont le GPT semble prendre en compte
Interactivité	GPP « Posez à chaque équipe une question tirée du questionnaire ou formulez vos propres questions. Laissez une minute aux élèves pour y répondre individuellement, par écrit. Ensuite, dites-leur d'en discuter ensemble pendant que vous dirigez vers une autre équipe » (GPP, 2010, p. 119).	Ces extraits en effet suggèrent le travail en équipe et des discussions, ce qui crée une certaine interaction entre les évalués.
	GPI « L'activité peut se faire en équipes de deux ou individuellement. <ul style="list-style-type: none"> • Allouer cinq minutes aux élèves pour qu'ils lisent l'activité, déterminent les photos qui montrent un travail s'accomplissant sur l'objet, et expliquent leur réponse à leur camarade. • Si les élèves font l'activité individuellement, discuter des réponses avec la classe » (GPI, 2009, p. 59). « Allouer cinq minutes pour qu'ils réalisent l'activité et qu'ils comparent leurs réponses avec celles d'une ou d'un camarade » (GPI, 2009, p. 37).	Dans cet extrait, on recommande éventuellement de faire travailler les élèves en équipe, ce qui crée une interactivité entre eux.
	GPT	(Extrait trop long pour être inséré ici).