



Planning, Research and Evaluation
Department of Education, Culture and Employment

Planification, recherche et évaluation
Ministère de l'Éducation, de la Culture et de la Formation

Strengthening Teacher Instructional Practice (STIP) Pilot Evaluation Plan

Renforcement des pratiques d'enseignement Plan d'évaluation pilote

Le présent document contient un résumé en français.

Table of Contents

Executive Summary	3
Résumé.....	4
1.0 Overview	5
2.0 Description of the Program: Strengthening Teacher Instructional Practices Pilot Project	7
2.1 Strengthening Teacher Instructional Practices Pilot Project Design.....	8
3.0 Evaluation Purpose	10
3.1 Evaluation Questions	10
4.0 Evaluation Design	12
4.1 Principles Underlying the Evaluation	12
4.2: STIP Logic Model.....	14
5.0 Stakeholder Engagement Strategy	15
6.0 Evaluation Methodology.....	16
6.1 Data Collection Methods	16
6.2 Evaluation Question Matrix with Indicator Descriptions and Data Sources and Tools.....	17
6.3 Data Collection Tools.....	21
6.4 Data Collection Protocols	23
6.5 Analysis Plan.....	23
7.0 Reporting Schedule	24
8.0 Workplan and Budget.....	25
8.1 Workplan	25
8.2 Budget	28
9.0 Limitations, Challenges, and Mitigation Strategy	28
Appendices:.....	30
Evaluation Tools.....	32

Executive Summary

The Strengthening Teacher Instructional Practice (STIP) pilot project is anticipated to have a positive impact on our teachers, especially in relation to:

- reduced teacher workloads;
- increased effectiveness of teacher professional collaboration and learning opportunities;
- improved teacher wellness; and
- improved quality of instruction.

Grounded in evaluation theory, this evaluation plan we will measure the extent to which these anticipated impacts on teachers have been realized by answering the following evaluation questions:

1. How have the hours of instruction been implemented and administered at participating schools?
2. To what degree has STIP supported reductions in teacher workloads?
3. To what degree has the STIP project supported quality professional collaboration and learning within NWT schools?
4. To what degree has the STIP project supported teacher wellness in NWT schools?
5. To what degree has the STIP project supported improvements in the quality of instruction in NWT schools?
6. What level of academic achievement are students experiencing at STIP participating schools; recognizing that student achievement cannot be directly correlated with the level of STIP success?
7. Were there any unexpected outcomes because of the STIP project?

Multiple data sources will be used to answer these questions, including various sets of administrative data and feedback from various stakeholders. Using such an approach will assure that the findings of the STIP pilot project are fully understood and valid. These findings will be reported through two deliverables: a formative evaluation undertaken midway through implementation of the pilot project (2018-2019) and a final, summative, evaluation after the pilot's conclusion (2020-2021).

Résumé

On s'attend à ce que le projet pilote « Renforcement des pratiques d'enseignement » ait un effet positif sur les enseignants, particulièrement en ce qui a trait à:

- la réduction de leur charge de travail;
- l'amélioration de l'efficacité du travail d'équipe et des possibilités d'apprentissage;
- l'amélioration de leur bien-être;
- l'amélioration de la qualité de l'enseignement.

Fondé sur la théorie de l'évaluation, le plan d'évaluation fait appel aux questions suivantes pour déterminer dans quelle mesure le projet a produit les bienfaits escomptés chez les enseignants:

1. Comment les heures d'enseignement ont-elles été mises en œuvre et administrées dans les écoles participantes?
2. Dans quelle mesure le projet pilote a-t-il facilité la réduction de la charge de travail des enseignants?
3. Dans quelle mesure le projet pilote a-t-il soutenu une collaboration et du perfectionnement professionnels de qualité dans les écoles des Territoires du Nord-Ouest (TNO)?
4. Dans quelle mesure le projet pilote a-t-il favorisé le bien-être des enseignants dans les écoles des TNO?
5. Dans quelle mesure le projet pilote a-t-il contribué à l'amélioration de la qualité de l'enseignement dans les écoles des TNO?
6. Quels sont les résultats scolaires des élèves dans les écoles qui participent au projet pilote (même si la réussite des élèves ne peut être directement liée à la réussite du projet)?
7. Le projet pilote a-t-il eu des résultats inattendus?

Plusieurs sources de données seront utilisées pour répondre à ces questions, y compris divers ensembles de données administratives et des commentaires formulés par divers intervenants. Cette approche éclairera et validera les conclusions du projet pilote « Renforcement des pratiques d'enseignement ». Celles-ci seront ensuite transmises de deux façons : une évaluation formative menée à mi-chemin dans la mise en œuvre du projet pilote (2018-2019) et une évaluation sommative définitive après la conclusion du projet (2020-2021).

1.0 Overview

Students in the Northwest Territories (NWT) have among the highest number of instructional hours in Canada (Figure 1). NWT instructional hours are well above the Canadian average, yet educational outcomes, including graduation and attendance rates, remain poor (Figures 2 and 3). Time for teachers to carry out non-instructional teacher duties during the school day is also not guaranteed in the NWT, unlike some Canadian jurisdictions. In addition to these challenges, the Government of the NWT (GNWT)'s Department of Education, Culture and Employment (ECE) has learned from its recent *NWT Teacher Time and Workload Study* (ECE, 2015) that many NWT teachers are working long hours and experience work-life balance challenges.

Figure 1: Instructional Hours Across Canada (Upper Secondary Schools 2015-2016)

Province	<i>Instructional Hours (early 2000s)¹</i>	<i>Instructional Hours (current)²</i>
British Columbia	963	952
Alberta	1000	1000
Saskatchewan	940	950
Manitoba	1045	1018
Ontario	950	880
Quebec	900	900
New Brunswick	1064	1018
Nova Scotia	930	935
Prince Edward Island	925	880
Newfoundland and Labrador	950	935
NWT	1045	1045 (current) 945 (new 2017-2018)

Figure 2: High School Graduation Rates (2013-2014)

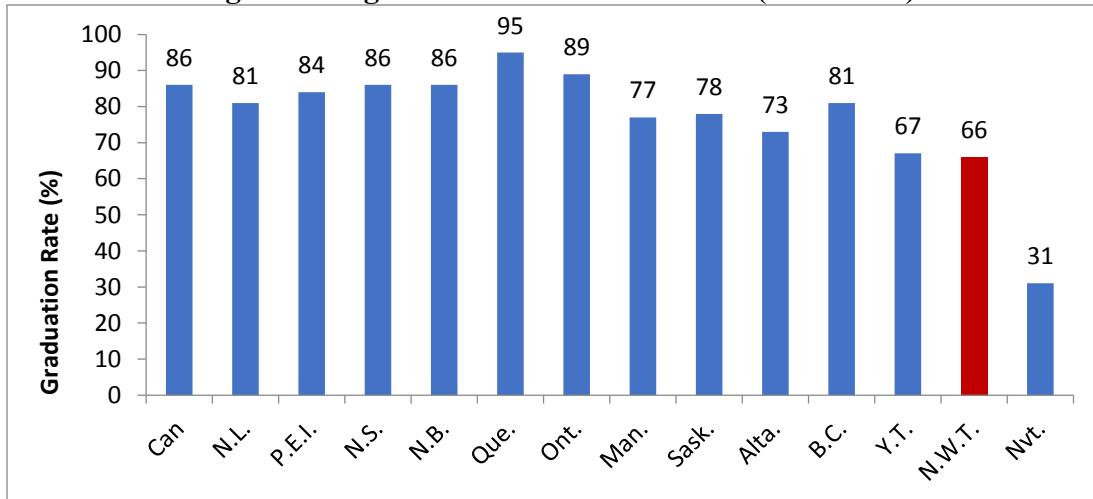
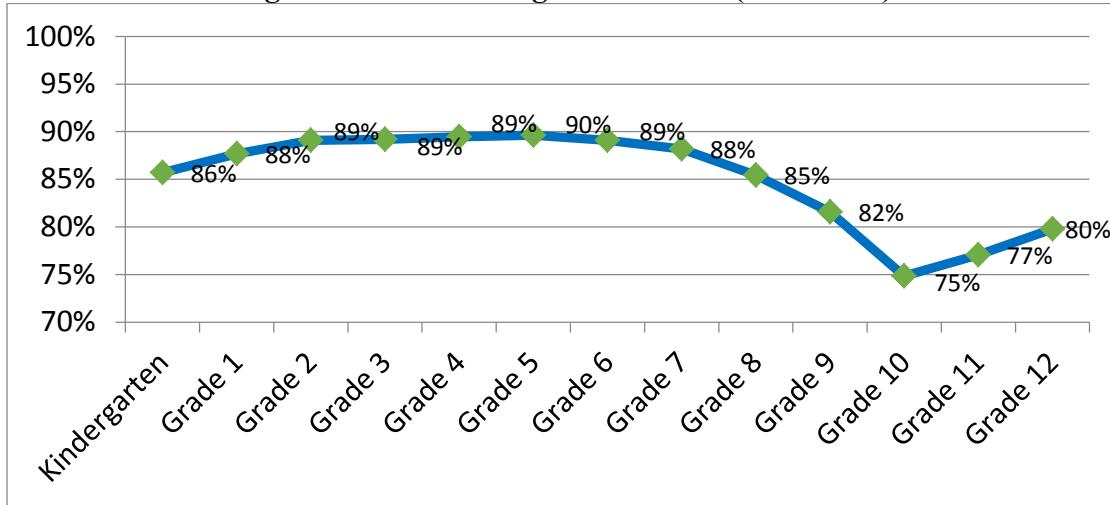


Figure 3: NWT Average Attendance (2007-2016)



Changes are needed to address these multiple and compounding challenges. ECE is responsible for supporting NWT school boards, NWT teachers, schools, and most importantly, students, so that all students can achieve their full potential. To this end, in 2013 ECE embarked on a series of broad changes to the NWT education system, collectively known as the Education Renewal and Innovation initiative (Education Renewal).

The work of Education Renewal was also prompted in part by the recommendations of the GNWT's Program Review Office and from the Office of the Auditor General (2010), as well as through growing evidence from NWT data, research and engagement with many education partners. While there are many examples of extraordinary schools, teachers and learning programs in the NWT, the data, research and feedback from partners indicate that the current approach to education is not producing the overall levels of student achievement that should reasonably be expected when compared to the rest of Canada from the investments that are being made. Education Renewal is a collection of many complementary projects and initiatives, designed to support NWT students and teachers, using evidence-based practices, traditional knowledge, and 21st century learning practices.

2.0 Description of the Program: Strengthening Teacher Instructional Practices Pilot Project

The Strengthening Teacher Instructional Practices (STIP) pilot project is one such initiative within Education Renewal, listed under Area for Action 10: Supporting Northern Professionals in the June 2015 *Three-Year Education Renewal Action Plan*. It was conceived following research and investigation into the following topics: NWT teacher workloads, the relationship between student achievement and instructional time, factors that have greatest impact on student learning, and best practices in teacher professional development.

Findings related to NWT teacher workloads came primarily from the *NWT Teacher Time and Workload Study* (ECE, 2015). This research study found that the total average weekly work time reported by respondents in this study (52.2 hours) corresponded with the 50-55 hours commonly reported in other Canadian teacher workload studies (BC Teachers' Federation, 2016). Key concerns mentioned in this study, as well as in other workload studies, include: the amount of time teachers spend outside of the regular instructional day on work tasks; classroom complexity (related to behavioural issues and student needs); and increased work demands, often related to reporting and non-instructional duties.

There is substantial literature on work intensification which argues that work demands for educators are growing without the accompaniment of more time, support, or resources for them to adequately carry out their work functions (Easthope & Easthope, 2000; Wotherspoon, 2008; Hargreaves, 1994). The increase in demands, in turn, is deemed to result in increased stress and work-life balance concerns (NWTTA, 2013; Naylor & White, 2009).

Findings related to the relationship between student achievement and instructional time came primarily from a literature review conducted by ECE. This literature review found several comprehensive studies and reports which investigated the relationship between *student achievement* and *instructional time*, providing evidence that there is no strong relationship between these variables (Long, 2014; Baker et al., 2004; Patall et al., 2010; Van Damme, 2014, Noonan, 2007; Hattie, 2015a). The main conclusion drawn in the research is that more attention should be given to improving *instructional quality*, such as teaching practice and curriculum, rather than *instructional quantity*. Thus, more instructional time does not mean better student achievement.

Research on factors that have the greatest impact on student learning revealed that *teachers* and the *teaching approaches* they employ, as opposed to *school*, *parents* and *student* factors, contribute most to student learning (Hattie, 2009). Furthermore, research on best practices in teacher professional development found that ongoing, job-embedded, professional development, such as professional learning communities, are most effective (Vescio et al., 2008; Villegas-Reimers, 2003; Hattie, 2015b), and that these forms of collaborative learning is associated with higher teacher efficacy and job satisfaction (TALIS, 2008).

2.1 Strengthening Teacher Instructional Practices Pilot Project Design

Given these challenges, the STIP pilot project is intended to create system- level changes to better support schools in offering structured opportunities for teachers to improve their instructional practice. Likewise, it is intended to recognize the importance for teachers to have time devoted within the school day to carry out tasks such as planning, assessing, and reporting.

The STIP pilot project is designed to offer schools the opportunity to opt into a pilot project, whereby the number of instructional hours for students can be re-directed by up to 100 hours per school year. These 100 hours must then be used instead for teacher time as follows:

- Up to 50% of any hours re-directed instructional time through STIP can be allocated to collaborative professional learning led by the School Principal and/or the Superintendent of Education. This could include participation communities of practice, peer teamwork, and professional learning communities.
- At least 50% of any hours re-directed through STIP must be allocated towards individual professional time for teachers. This is time teachers can use for planning, assessing, reporting and individual professional learning.

The expected benefits of fewer classroom instruction hours will first accrue to teachers, who are anticipated to experience improvements in their teaching practice and wellness. Teaching practice can be expected to improve as a result of participation in structured, collaborative professional learning opportunities, as well as having embedded time for lesson planning and professional duties. An increase in teacher wellness can also be expected as teachers will have more job-embedded time to complete work tasks, such as planning and assessment, which might otherwise be completed in their personal time and contribute to poor work-life balance. Participation in collaborative learning opportunities may also result in a greater collaborative culture and peer-to-peer support within schools, which should also have a positive impact on teacher wellness.

It is further anticipated that secondary benefits will then accrue to students, who will experience a higher quality of teaching. Together with other initiatives and projects undertaken through the Education Renewal and Innovation Framework, this will support students to achieve their full academic potential. The expected outcomes of STIP activities are outlined in the STIP Logic Model, found in Table 3.

STIP will be in place as a pilot project for three years (from the 2017-2018 school year, to the 2019-2020 school year). Schools have the option of participating for one, two, or all three years. There have been no limits on the number of NWT schools who may participate. The only criteria schools must meet to participate are as follows:

- maintain at least 945 instructional hours per year.

The mandatory components of how the re-directed time must be used include:

- individual professional duties and learning (such as lesson planning and reporting of student progress, meeting with parents and families of students and developing student support or individual education plans and learning); and
- collaborative professional learning led by the School Principal and/or the Superintendent of Education.

Schools must reapply through their Superintendent to the STIP Committee to participate each year, to ensure they are meeting the STIP project criteria. The STIP Committee is made up of ECE, NWT Teachers Association (NWTTA) and the NWT Superintendents Association (NWTSA). More information on this committee and other stakeholder groups in the evaluation is found in section 5, below. Following the conclusion of the pilot project, the results of the evaluation will be used to make changes to STIP by, in the least, recommending best practices and models for STIP implementation.

3.0 Evaluation Purpose

Changes to any aspect of the education system will provoke concern for how they will affect student outcomes, and for unexpected outcomes that may result from those changes. Education Renewal has committed to undertaking evaluation on pilot projects and initiatives as they unfold to ensure that changes are best supporting students and teachers.

This evaluation is designed to ensure that Members of the Legislative Assembly (MLAs), ECE, the NWT TA, as well as NWT parents and students have the information and evidence they need to determine whether STIP is being implemented as it was intended, whether improvements could be made to the project design and implementation, and whether, after the three-year pilot-project term has expired, it is supporting improved teacher wellness and quality instruction. This information will support ECE in determining the extent to which the project should be continued.

The evaluation will gather this evidence through ongoing, annual data collection and analysis, and present it to the public at large through formal reporting, culminating in a final summative evaluation within six months following the conclusion of the 2019-2020 academic year that includes all three years of pilot project results and findings. The reports of this evaluation will be tabled by the Minister at the sitting of the Legislative Assembly during which the review is completed; or at the next sitting of the Legislative Assembly, if the Legislative Assembly is not sitting when the report is completed.

3.1 Evaluation Questions

The evaluation questions are based on the STIP logic model in Table 3. Because the logic model moves from immediate effects on the left, towards short, medium and long term effects on the right, not all evaluation questions can be answered for each year of the evaluation. The evaluation questions will guide the information provided in the evaluation report(s). Not all stakeholders will be expected to provide input on each evaluation question.

The first evaluation question is intended to examine the hours of instruction in STIP participating schools:

1. How have the hours of instruction been implemented and administered at participating schools?
 - This question looks at the input and output level information that is outlined in the logic model

Evaluation questions 2 to 3 are intended to examine how effective the administration and implementation of STIP hours have been in terms of supporting teacher workloads and quality professional collaboration and learning:

2. To what degree has STIP supported reductions in teacher workloads?
 - This question looks at short term outcomes as outlined in the logic model
3. To what degree has the STIP project supported quality professional collaboration and learning within NWT schools?
 - This question looks at short term outcomes as outlined in the logic model

Evaluation questions 5 to 6 are intended to examine the degree to which the objectives of the STIP pilot project were achieved:

4. To what degree has the STIP project supported teacher wellness in NWT schools?
 - This question looks at medium-term outcomes as outlined in the logic model
5. To what degree has the STIP project supported improvements in the quality of instruction in NWT schools?
 - This question looks at the medium-term outcomes outlined in the logic model
6. What level of academic achievement are students experiencing at STIP participating schools?
 - This question looks at the long-term outcomes as outlined in the logic model

The final evaluation question is intended to capture information related to unexpected outcomes of the STIP pilot project:

7. Were there any unexpected outcomes because of the STIP project?
 - This question is not based on the logic model, but instead seeks to gather information on other effects STIP may have on students, teachers, and families

4.0 Evaluation Design

This evaluation uses a mixed methods approach¹ which means that qualitative information (gathered from sources such as interviews, focus groups, and survey responses from teachers, parents, and students, etc.) is collected alongside quantitative data (gathered from sources such as human resources information systems, student records systems, surveys, etc.) and jointly analyzed to produce comprehensive findings that are grounded in a triangulation of evidence.

The evaluation will take place over the course of three years, beginning in 2017. Data collection and analysis will take place throughout the three years, with a formative evaluation reports completed in 2018-2019, and a final summative evaluation and associated report completed after the three years of data collection are done. The schedule of evaluation reports can be found in Table 7: Reporting Schedule.

4.1 Principles Underlying the Evaluation

Guskey's framework for evaluation of professional development outlines five levels of consideration when evaluating the effectiveness of professional development in an education setting. It stresses the importance of improved student outcomes as the goal of any professional development activity, making it an appropriate framework through which to evaluate the STIP pilot project, which shares this long-term outcome of improved student outcomes as part of Education Renewal. Furthermore, the emphasis placed on supporting professional development through STIP makes this framework an appropriate choice for the evaluation by helping to break down and better understand the different types and levels of effectiveness of professional development activities undertaken through STIP.

In addition to student outcomes, which will be discussed further below, Guskey's framework suggests looking at: participant reactions to professional development; participant learning in professional development activities; organizational support for professional development; and participant use of new skills learned through professional development (Guskey, 2002). Appendix A shows how these domains will be captured in the STIP evaluation.

A learning evaluation is an approach to qualitative and quantitative data collection and analysis in interventions with multiple organizations where continuous program improvement is a goal. It blends real time assessments, shared indicators, and rapid quality improvement feedback mechanisms to encourage learning across organizations. In a learning evaluation, continued communication of results to the implementation group is a central feature. Balasubramanian et al. (2015) suggest five principles for conducting a learning evaluation. These are laid out in the table below, together with strategies that ECE and NWT schools can use to address the principles.

This is an appropriate complement to Guskey's framework because of its focus on ongoing learning and program improvement, and because of the framework's treatment of multiple organizations. Because STIP involves many different schools who are each implementing the STIP project in different ways,

¹ In using a mixed-methods approach, this evaluation is grounded in two evaluation frameworks; Guskey's (2002) framework for evaluation of professional development, and Balasubramanian et al's (2015) learning evaluation framework.

using a different number of redirected hours, and towards a different breakdown between time for professional duties and learning as well as professional collaborative learning, it is important for the evaluation to be able to work with variations to the program design.

In part, this will be accomplished by creating grade-based categories within the participating schools so that the unit of analysis is on elementary and junior high schools (Junior Kindergarten to grade 9) and high schools (grades 10 to 12). These grade-based categories were chosen both for practicality, so that the large number of participating schools can be organized into manageable clusters from which to draw results, and due to the different types of concerns for the potential unintended consequences of STIP that could occur across the different grade-based categories.

The learning evaluation approach will also enable ECE to examine the context of the different schools participating in STIP and ensure that variations across different school types are accounted for. Appendix B demonstrates how the principles of a learning evaluation will be followed in the STIP evaluation.

4.2: STIP Logic Model

A logic model has been designed for this pilot project to help articulate the purpose of the pilot project, to help inform the key activities and outcomes that will be monitored for the evaluation, and to show how the changes to instructional hours are anticipated to contribute to student outcomes².

Table 1: STIP Logic Model

Inputs	Outputs Activities	Short	Medium	Long / Impact
<p>Re-development of STIP-participating school calendars to reflect a 1-100 hour redirection of instructional hours by NWT education bodies</p> <p>Support for redirected instructional hours by NWT school bodies and teachers, through development of proposals</p> <p>Support for redirected instructional hours by STIP Committee through proposal review</p>	<p>Up to 50% of redirected hours are reallocated towards professional duties and learning for teachers</p> <p>Up to 50% of redirected hours are reallocated towards collaborative professional learning for teachers</p>	<p>Reduced teacher workloads at participating NWT schools</p> <p>Increased effectiveness of teacher professional collaboration and learning opportunities at participating NWT schools</p>	<p>Improved teacher wellness at participating NWT schools</p> <p>Improved quality of instruction at participating NWT schools</p>	<p>Improved opportunity for student achievement at participating NWT schools</p>
<p>Assumptions</p> <ul style="list-style-type: none"> -Student achievement is supported by classroom instructional hours; however, student achievement plateaus after a base range of instructional time, after which point student achievement does not increase with an increase in instructional time. -Approximately 945 hours of instructional time is a suitable base range, after which point an increase in instructional hours will not greatly contribute to NWT student achievement - Improving the quality of teaching is more effective at supporting student achievement than increasing instructional hours (once the 945 hour base threshold is met) -There is room for improving the quality of teaching in the NWT - NWT teachers are willing and able to work on their classroom instructional skills - Teacher wellness contributes to teaching quality - Collective teacher efficacy is one of the most important factors in terms of improving student achievement 				
<p>External Factors</p> <ul style="list-style-type: none"> - Quality of professional learning opportunities - Fit between professional learning opportunities and teacher needs - Teacher and student motivation to support student achievement - Family and community support for student achievement - Additional, co-occurring and concurrent programs and initiatives designed to support student success and achievement in NWT schools, including (but not limited to), the Alternative High School Pathways initiatives, Distance Learning initiatives, I Am More Than ABC's campaign, Anti-Bullying and other student wellness initiatives, etc. 				

² Student achievement is a long term outcome as per the logic model in Table 1 and cannot be directly correlated with the level of STIP success. The level of STIP success can only contribute to the opportunity for improved student achievement.

5.0 Stakeholder Engagement Strategy

The Minister of ECE is ultimately responsible for ensuring that STIP is evaluated, and that the evaluation provides NWT teachers, parents and guardians, students, MLAs, and members of the public with useful information that can lead to informed decision making about the STIP pilot project.

While all NWT residents have an interest and stake in the STIP pilot project and this evaluation, the following primary stakeholders have been identified:

- NWT teachers working in STIP-participating schools
- NWT school administration in STIP-participating schools
- NWT students in STIP-participating schools
- NWT parents and guardians of students in STIP-participating schools

In addition to these primary stakeholders who are being impacted by STIP, the following secondary stakeholder groups have decision making responsibilities related to STIP:

- The Department of Education, Culture, and Employment
- Members of the Legislative Assembly
- NWT Teachers Association
- NWT Superintendents Association

In order to fully engage the primary stakeholders, an **Evaluation Steering Committee** will be struck to guide the STIP evaluation. The members of the Evaluation Steering Committee will be responsible for ensuring their unique stakeholder perspectives are represented in the evaluation. They also provide oversight and high-level management of the STIP evaluation. The committee will help to interpret and validate analysis conducted by the Evaluation Working Group (described below). The Evaluation Steering Committee will be composed of the following representatives:

- The President of the NWT Teachers' Association
- Designate from the NWT Superintendents Association
- The Education and Culture and Corporate Services Assistant Deputy Ministers of the Department of Education, Culture, and Employment

The **Evaluation Working Group** will be struck to carry out the evaluation plan as described in this document. The members of this Working Group will be responsible for ensuring the evaluation takes place and will be composed of:

- Directors and related staff of the following divisions within ECE:
 - Planning, Research and Evaluation
 - Education Operations
 - Teaching and Learning
 - Health, Wellness and Student Support
- Superintendents of NWT Education Bodies
- Executive Director and Assistant Executive Director of the NWT Teachers Association

A Terms of Reference will be established for both the Evaluation Steering Committee and the Evaluation Working Group.

The evaluation plan was developed for the Steering Committee by the Department of ECE's Planning, Research, and Evaluation (PRE) Division, and will be the unit responsible for implementing the Plan once approved. Through the development of the evaluation plan and the implementation of the evaluation, PRE is responsible for ensuring the evaluation design, analysis and reporting are methodologically sound, and follows evaluation standards, ethics, and best practices. Primary stakeholders will also be engaged through data collection. Each primary stakeholder group will be surveyed through annual stakeholder surveys (see Tables 5 and 6), and input from these groups will be collated and included in the evaluation reports.

6.0 Evaluation Methodology

As noted, the evaluation follows a mixed methods approach, grounded in the two evaluation frameworks described above.

All data will be collected through several data collection methods and tools and analyzed according to school type and relevant evaluation question. Each of these elements are described in more detail in an evaluation matrix presented in the corresponding section below. The proposed evaluation matrix will serve as a methodological guideline for the evaluation; however, given the length and complexity of the pilot project, adjustments to outlined indicators and specific data collection methods and tools may need to be made in the course of the evaluation.

6.1 Data Collection Methods

Most of the STIP evaluation data collection methods and tools will need to be specifically developed for the purposes of this evaluation; however, there are some existing sources of data that ECE can draw on. This includes the proposals schools must submit to be part of the STIP pilot project, and human resource data such as the number of sick days and non-vacation leave that NWT teachers take.

ECE has also begun collecting stakeholder feedback through its Accountability Framework in 2017. The 2017 NWT Classroom Teacher Survey is directed at NWT teachers, and includes many STIP related questions that could be used for comparison purposes in the STIP evaluation reports.

Though not reported on as part of this evaluation, student outcome data is also already captured by ECE through tools including (but not limited to) the Early Development Instrument, Middle Years Development Instrument, the Health Behaviour of School Age Children Survey, Alberta Student Achievement testing results, and the PowerSchool system which tracks attendance and high school graduation rates. ECE will continue to collect, monitor, and report on student outcome data through the Department's Accountability Framework reports.

6.2 Evaluation Question Matrix with Indicator Descriptions and Data Sources and Tools

Table 2: Evaluation Matrix

Evaluation Question	Indicator	Data Source	Data Collection Method or Tool
Hours of instruction in STIP participating schools:			
1. How have the hours of instruction been implemented and administered at participating schools?	a. consistency between allocation of hours proposed, and actual allocation of redirected teaching hours	1A1. STIP project proposals, including school calendars 1A2. NWT teachers and administration at participating STIP schools	1A1. Proposals that are collected prior to each school year for participation in STIP. 1A2. Annual teacher and administration survey
Effectiveness of the administration and implementation of STIP hours:			
2. To what degree has STIP supported reductions in teacher workloads?	a. teacher reported increase (decrease) in time spent planning for classes during work hours b. teacher reported decrease (increase) in time spent planning for classes after work hours c. administration reported increase (decrease) in time made available to teachers during work hours	2A1. Follow up Teacher Workload Study 2B1. Follow up Teacher Workload Study 2C1. School administration at participating STIP schools	2A1. Teacher Workload Study questionnaire 2B1. Teacher Workload Study questionnaire 2C1. Annual STIP proposals

Evaluation Question	Indicator	Data Source	Data Collection Method or Tool
3. To what degree has the STIP project increased the opportunity for professional collaboration and learning for teachers within NWT schools?	<ul style="list-style-type: none"> a. # and type of professional collaboration and learning opportunities offered through participating STIP schools b. rate of participation in professional collaboration and learning opportunities offered at participating STIP schools c. teacher reported satisfaction with professional collaboration and learning opportunities d. administration reported satisfaction with professional collaboration and learning opportunities e. Teacher learning outcomes of professional collaborative learning and individual learning opportunities 	<ul style="list-style-type: none"> 3A1. NWT school boards 3B1. NWT school boards 3C1. NWT teachers at participating STIP schools 3D1. NWT administration at participating STIP schools 3E1. Results of pre and post-testing from professional collaborative learning and individual learning opportunities, where available 	<ul style="list-style-type: none"> 3A1. NWT school board records 3B1. NWT schools board records 3C1. Annual teacher survey 3D1. Annual administration survey 3E1. Pre- and post-workshop competency testing
Degree to which the objectives of the STIP pilot project were achieved			

4. To what degree has the STIP project supported teacher wellness	<ul style="list-style-type: none"> a. teacher reported self-assessment of work-life balance and work stress b. administration reported assessment of teacher work-life balance at their school c. # of days of sick/disability/etc. Leave taken by NWT teachers in 2017/18, 2018/19 and 2019/20 compared to 2016/17 (or in STIP compared to non-STIP schools) 	<p>4A1. NWT teachers at participating STIP schools</p> <p>4A2. Follow up Teacher Workload Study</p> <p>4B1. NWT administration at participating STIP schools</p> <p>4C1. PeopleSoft Human Resource Management System (Department of Human Resources) and YK1/YCS Human Resources</p>	<p>4A1. Annual teacher survey</p> <p>4A2. Follow up Teacher Workload Study questionnaire</p> <p>4B1. Annual administration survey</p> <p>4C1. PeopleSoft Human Resource Management System (Department of Human Resources)</p>
Evaluation Question	Indicator	Data Source	Data Collection Method or Tool
5. To what degree has the STIP project supported quality instruction in NWT schools?	<ul style="list-style-type: none"> a. teacher reported self-assessment of teacher efficacy b. administration reported assessment of teaching quality c. student reported assessment of teaching quality 	<p>5A1. NWT teachers at participating STIP schools</p> <p>5B1. NWT administration at participating STIP schools</p> <p>5C1. NWT students at participating STIP schools</p>	<p>5A1. Annual teacher survey</p> <p>5B1. Annual administration survey</p> <p>5C1. Annual student survey</p>
6. What level of academic achievement are students experiencing at STIP participating schools?	<ul style="list-style-type: none"> a. AAT scores b. Graduation Rates c. Attendance 	<p>6A1. ECE aggregated student records</p> <p>6B1. ECE aggregated student records</p> <p>6C1. ECE aggregated student records</p>	<p>6A1. AAT exams already administered in schools</p> <p>6B1. NWT Student Records</p> <p>6C1. NWT Student Records</p>
Degree to which unexpected outcomes were experienced:			

<p>7. Are parents, students, or teachers experiencing unexpected outcomes of the STIP project?</p>	<p>a. Teacher reported 'other' outcomes of STIP b. Administration reported 'other' outcomes of STIP c. Parent reported 'other' outcomes of STIP d. Student reported 'other' outcomes of STIP</p>	<p>6A1. NWT teachers at participating STIP schools 6B1. NWT administration at participating STIP schools 6C1. NWT parents at participating STIP schools 6D1. NWT students at participating STIP schools</p>	<p>6A1. Annual teacher survey 6B1. Annual administration survey 6C1. Annual parent survey 6D1. Annual student survey</p>
--	---	--	---

6.3 Data Collection Tools

The data for this evaluation will come from key stakeholders, including: teachers; school administration; parents or guardians of students; and in some cases, students themselves, all of whom will belong to one of the STIP-participating schools. Information from these stakeholders will be gathered annually via stakeholder surveys, tailored to each of the groups, and in some cases adapted from year to year.

Stakeholder surveys will be administered through multiple avenues to increase response rates and survey reach. Table 4 describes these tools and distribution mechanisms avenues.

Table 3: Annual stakeholder survey tool descriptions

Stakeholder group being surveyed	Survey administration mechanism	Relevant Evaluation Question and Indicator*
Teachers at STIP-participating schools	- NWT Classroom Teacher Survey	#2a, #2b, #3c, #4a, #5a, #6a
Schools administration staff at STIP-participating schools	- School Administrator Survey	#2c, #3d, #4b,
Parents/guardians with students at STIP-participating schools	- Online survey distributed to parents/guardians by STIP participating schools through ECE's website and social media platforms, including Facebook and Twitter; and - Paper-based survey distributed by STIP-participating schools	#5c, #6b
High school students at STIP-participating high schools	- Online survey distributed to high school students by STIP participating schools through ECE's website and social media platforms, including Facebook and Twitter; and - Paper-based survey distributed by STIP-participating schools	#5b, #6c

*See Table 4 for indicator descriptions

In addition to stakeholder surveys, the evaluation relies on data collected through existing systems and administrative processes. Table 5 describes these existing systems, administrative processes, and the evaluation questions and indicators to which they are related.

Table 6: Existing administrative processes and systems used for the evaluation data collection

Existing administrative process or system	Type of data extracted for use in the STIP evaluation	Relevant evaluation question and indicator*
Human Resource Information System (HRIS) / PeopleSoft	<ul style="list-style-type: none"> - Aggregated data regarding teacher use of leave time 	#4c
STIP proposals, submitted by each STIP-participating school in the year prior to participation	<ul style="list-style-type: none"> - Hours of planned STIP time - Percentage breakdown of STIP hours 	#1a, #2c,
STIP-participating school records related to collaborative professional learning opportunities	<ul style="list-style-type: none"> - Collaborative professional learning opportunities offered through STIP hours - Participation rates in professional collaborative learning opportunities offered during STIP hours 	#3a, #3b
Student achievement data ³	<ul style="list-style-type: none"> - Graduation Rates - AAT scores - Attendance Rates 	#7

*See Table 4 for indicator descriptions

Two final data collection tools are required for this evaluation. Both are directed at teachers in STIP-participating schools, and are important for capturing levels of collaborative professional learning according to the evaluation principles listed in section 4.1 above. These are: professional learning satisfaction tools; and, pre- and post-professional collaboration and learning assessments. They will be used towards evaluation questions and indicators 3c and 3e.

It is critical that the STIP-participating school administration work with the Evaluation Working Group to create these tools so that they can be administered before and after professional collaboration and learning opportunities; likely at the beginning and end of the school year as these opportunities are intended to be embedded and ongoing.

³ Student achievement is a long term outcome as per the logic model in Table 1 and cannot be directly correlated with the level of STIP success. The level of STIP success can only contribute to the opportunity for improved student achievement.

6.4 Data Collection Protocols

All data collected through the evaluation will be done following the GNWT's *Access To Information and Privacy Act* (ATIPP Act).

Unique data collected and reported through STIP will be saved for 7 years upon tabling of the summative evaluation report in the Legislative Assembly and then transferred to the custody of the Northwest Territories Archives, where the records will either be selected for preservation as a historical record or appraised as having no historical value and destroyed. Data collected through existing ECE or GNWT data collection tools and mechanisms will follow the data collection protocols related to that system or process.

All provision of data by individual stakeholders is voluntary and will be collated to assure results will not be personally identifiable. This will be explicitly communicated to stakeholders directly.

6.5 Analysis Plan

For qualitative data, thorough content analysis will be the primary method of analysis. Survey and focus group responses will be reviewed by the Evaluation Working Group and coded for key themes. It is anticipated that the Evaluation Steering Committee will play an active role in helping to interpret themes and validate the analysis conducted by the Evaluation Working Group.

For quantitative data that is not already collected through Accountability Framework reporting, the evaluators will work with system administrators to determine the best point in time to extract pre-identified data-sets from the corresponding data-system to ensure it is available for the formative and summative evaluation reports.

Once gathered, both qualitative and quantitative data will be used together to answer the appropriate questions as per the Evaluation Matrix (Table 6). All analysis will be reviewed by the STIP Evaluation Steering Committee. The reports of this evaluation will be tabled by the Minister at the sitting of the Legislative Assembly during which the review is completed; or at the next sitting of the Legislative Assembly, if the Legislative Assembly is not sitting when the report is completed.

Data will be analyzed according to the school grade types identified earlier: elementary school (junior kindergarten to grade 9); and high school (grade 10 to 12). These were chosen as categories of analysis because it is anticipated that findings for each of the evaluation questions may differ across school grade types. Evaluation reports will present findings for each evaluation question broken down by these school types.

7.0 Reporting Schedule

This evaluation is designed to produce two reports, a formative report covering the implementation and early findings of the first two years of the pilot project, and a summative report after the final year that encompasses results from all three years. Both reports will be made public. This reporting schedule is described in the table below.

Table 4: Reporting Schedule

	Formative Evaluation Report	Final Summative Evaluation Report
Target release date	May 2019	August 2020
Data period included	2017-2019	2017-2020
Evaluation questions included	1	1-6

8.0 Workplan and Budget

The evaluation plan will be managed by ECE's PRE Division, with oversight by the Evaluation Steering Committee. Evaluation reports are published by the Department of ECE. PRE is responsible for ensuring the evaluation takes place according to the plans laid out in this document, and by the dates laid out in the Gantt charts in the Workplan section below. Any changes in resources, inputs, or other contextual factors related to STIP and the evaluation plan will result in changes to the workplan and should be updated accordingly.

PRE is also responsible for managing the budget associated with the STIP evaluation. The budget for this evaluation is laid out in the Budget section below. The budget itself comes from the EO Division of ECE. Any changes to the budget will result in changes to the workplan.

8.1 Workplan

Table 5: Pre-STIP Gantt Chart

		Pre-STIP		
		Jun 2017	Jul 2017	Aug 2017
Evaluation Plan Finalized				
Year 1 Data Collection Tools Developed				

Table 6: STIP Year 1 Gantt Chart

	STIP Year 1											
	Sep 2017	Oct 2017	Nov 2017	Dec 2017	Jan 2018	Feb 2018	Mar 2018	Apr 2018	May 2018	Jun 2018	Jul 2018	Aug 2018
Year 1 data collection tools finalized												
Year 1 data collection tools administered												

Table 7: STIP Year 2 Gantt Chart

	STIP Year 2											
	Sep 2018	Oct 2018	Nov 2018	Dec 2018	Jan 2019	Feb 2019	Mar 2019	Apr 2019	May 2019	Jun 2019	Jul 2019	Aug 2019
Year 2 data collection tools finalized												
Year 2 data collection tools administered												
Year 1 and 2 data analyzed and formative report drafted												
Final Year 2 Formative Report released												

Table 8: STIP Year 3 Gantt Chart

	STIP Year 3											
	Sep 2019	Oct 2019	Nov 2019	Dec 2019	Jan 2020	Feb 2020	Mar 2020	Apr 2020	May 2020	Jun 2020	Jul 2020	Aug 2020
Year 3 data collection tools finalized												
Year 3 data collection tools administered												
Years 1- 3 data analyzed and summative reported drafted												
Final summative evaluation report released												

8.2 Budget

A budget will be developed once the methodology is approved and actual costs can be determined.

9.0 Limitations, Challenges, and Mitigation Strategy

All evaluation plans face limitations and challenges. It is important that these are recognized and mitigated to the extent possible. All known limitations and challenges are identified below, along with corresponding mitigation strategies. Where additional limitations or challenges arise during the evaluation, an associated mitigation strategy will also be identified and added to this section.

Table 9: Limitations, challenges and mitigation strategies

Limitation or challenge	Mitigation Strategy
Perceived lack of stakeholder engagement and consultation in pilot project design among two primary stakeholder groups; parents/guardians, and students.	<p>Parents/guardians and students will have the opportunity to provide first hand input into the STIP evaluation, likely through survey feedback. Survey responses are one of the key data sources of this evaluation.</p> <p>Finally, parents/guardians and students, like all NWT residents, will have the opportunity to view final evaluation reports.</p>
High level of flexibility in how STIP is implemented across schools, resulting in high variation of pilot project design	<p>The evaluation methods and framework were chosen because they accommodate variations in project implementation. The frameworks and methods identify common aspects of STIP across schools.</p> <p>Breaking down STIP schools by school type (JK- grade 9; grades 10-12) for analysis purposes will help the Evaluation Working Group focus on results that are important to stakeholders and identify linkages to other aspects of Education Renewal.</p>
High level of stakeholder, including parent/guardian and MLA concern regarding potential for unexpected outcomes because of STIP	<p>Breaking down STIP schools by school type (JK- grade 9; grades 10-12) for analysis purposes will help the Evaluation Working Group focus on results that are important to stakeholders.</p>
Long term outcomes of STIP are influenced by a variety of projects and programs	<p>The issue of measuring and reporting on long-term outcomes that extend beyond the timeframe of an evaluation schedule is a common issue in evaluation reporting.</p> <p>To address this issue in the STIP evaluation, the Evaluation Working Group will rely on research that has established the link between the medium and long term outcomes of the STIP logic model. It will focus on confirming whether the STIP pilot project influences medium term outcomes.</p>
Influence of external factors on long term outcomes of STIP	<p>Additionally, while the opportunity for student achievement is identified as a long term STIP outcome, measurements of student outcomes are already monitored through existing ECE processes such as the Accountability Framework. Upon the conclusion of the STIP</p>

	<p>pilot, ECE will continue to track student outcomes and publicly report on these important indicators of the education system, however, the level of STIP success can only contribute to the opportunity for improved student achievement. Direct linkages cannot be made.</p>
Development of evaluation plan after pilot project design was already complete, including key data collection tools such as STIP proposals for Year 1	<p>The Evaluation Working Group will work with the information that is available in 2017 project proposals. This information has been reviewed and is suitable for satisfying the year 1 data collection needs.</p> <p>The Evaluation Working Group will coordinate more closely to support development of STIP proposal templates to ensure that any additional data requirements are satisfied in future STIP years</p>
Development of evaluation plan after development and launch of Accountability Framework 2017 Stakeholder Survey	<p>The Evaluation Working Group will work with the information that is available in the 2017 Accountability Framework Stakeholder Survey. This information has been reviewed and is suitable for satisfying the year 1 data collection needs.</p> <p>The Evaluation Working Group will coordinate more closely to support development of future Accountability Framework Stakeholder Surveys to ensure that any additional data requirements are satisfied in future STIP years</p>
Development of evaluation plan after the pilot project has been passed into legislation by NWT Legislative Assembly	<p>The Evaluation Working Group will ensure compliance with all current and new legislation as it prepares for, and implements, the STIP evaluation. The evaluation plan has been modified to capture necessary elements as described in NWT legislation regarding Instructional Hours.</p>

Appendices:

Appendix A:

Table 1: Framework for evaluating professional development and related STIP evaluation strategies

Levels of consideration in evaluating professional development	Related STIP evaluation strategy
Participant reactions to professional development	Collect information related to teachers' satisfaction with professional collaboration and learning opportunities within STIP-participating schools by administering satisfaction questionnaires to teachers. For 2017, this information is asked through the NWT Classroom Teacher's Survey.
Participant learning in professional development activities	Collect information related to how much teachers learn from professional collaboration and learning opportunities offered through STIP-participating schools through the NWT Classroom Teacher's Survey.
Organizational support for professional development	Collect information related to how school administrations select/offer professional collaboration and learning opportunities for teachers in STIP-participating schools, how closely STIP-schools follow the plans they submitted in their STIP proposals, and information regarding why there is variation, where variations from the plans occur.
Participant use of new skills learned	Collect information related to how teachers use the new skills they have learned from professional collaboration and learning opportunities offered through STIP-participating schools, by administering self-assessment tools to teachers.
Student outcomes	Together with all the related and overlapping pilot projects and initiatives being undertaken through the broader Education Renewal strategy, improving student outcomes is a long-term outcome of STIP as a pilot project, ECE will continue to monitor grade-appropriate student outcomes throughout the duration of the pilot project, and after. This information can be found in Education Body Operating Plans and Annual Reports. However, this evaluation will not draw independent or causal links between its activities and student outcomes for two reasons: <ol style="list-style-type: none">1. Student outcomes are impacted by a variety of complex, interacting factors beyond the scope and influence of STIP2. Teacher wellness and quality instructional practice are the key areas of focus for this pilot project

Appendix B:

Table1: Principles of a Learning Evaluation and related STIP evaluation strategy

Principles for Conducting a Learning Evaluation ⁴	Related Strategy
Principle 1: Gather data to describe types of changes made by organizations, how changes are implemented, and the evaluation of the change process.	Gather implementation results in years one and two, and use these findings to support schools who may be new applicants in years' two or three, or schools who wish to revise and tweak their approach in subsequent years.
Principle 2: Collect process and outcomes data that are relevant to organizations and to the research team.	Work directly with NWT schools through representation of the NWT TA and NWT Education Bodies on the Evaluation Steering Committee, as well as with the Legislative Assembly, to ensure that all indicators and findings are useful to ECE, NWT schools and MLAs.
Principle 3: Assess multi-level contextual factors that affect implementation, process, outcome, and transportability.	Collect information on contextual factors including regional context, grade-specific categories and results, and variations in STIP implementation and ensure these are communicated in evaluation findings.
Principle 4: Assist organizations in applying data to monitor the change process and make further improvements.	Through the Evaluation Steering Committee, ensure that results are communicated in a user-friendly format and timely fashion so that they can inform planning and decision making.
Principle 5: Operationalize common measurement and assessment strategies with the aim of generating transportable results.	Develop shared measures across schools and grade-categories that are relevant at a school level, a grade-category level, and a collective all-school, all-grade level.

The learning evaluation design will help ECE support continuous improvement within the STIP pilot project. Annual reporting on the results of each year of activity is part of the commitment to ongoing monitoring and rapid feedback. ECE has three years in which to achieve significant outcomes. The limited duration and funding makes rapid program improvement particularly important.

⁴Adapted from: Balasubramanian et al. Learning Evaluation: blending quality improvement and implementation research methods to study healthcare innovations *Implementation Science* 2015, 10:31

Evaluation Tools

The following data collection tools will have to be developed:

- Annual teacher survey (annual revisions)
- Annual administration survey
- Annual parent survey
- Annual student survey (for students in high school)
- Modifiable pre- and post- professional collaboration and learning opportunity assessment template

Changes to the evaluation methodology may prompt development of other data collection tools in addition or instead of those listed above.