### Project Background

#### 1. Description of Project

Complex competencies are a series of skills, capabilities, attributes and values that learners are expected to develop no matter what stages learners are in. It's not sufficient for current metrics to evaluate if young people are prepared to contribute to workplaces and communities. More metrics are necessary to demonstrate young people or students are equipped with a set of complex competencies. Ruby is such a tool that designed to support the implementation of competency-based, evidence-oriented, standards-references developmental assessment of complex competencies in teaching and learning of complex competencies. Ruby is thoughtfully designed to accommodate a wide array of users, including System Curriculum Leaders, Module Designers, Teachers, Students, School Principals, External Assessors, and Stakeholders or Parents.

Depending on evaluation objectives such as teacher planning, school reporting, or credentialing, Ruby dynamically tailors the array of multiple-choice questions. These questions are designed to identify and evaluate typical behavioral signs related to the competency under assessment. For example, an evaluation aimed at teacher planning may contain about 10 questions since its purpose is to assist a teacher in gaining a quick sense of a student's competencies for planning future lessons. Meanwhile, a credential-based assessment may consist of around 22 questions as it serves a more formal capacity in substantiating a student's skill set in the form of a certification.

Ultimately, by teachers' responses to the questions, Ruby platform generates different assessments to serve three main purposes:

**Teacher Planning**: Generates 'rocket reports' from individual assessments for lesson planning and promoting reflection among teachers and students.

**School Reporting**: Produces 'fan profiles' from aggregated assessments for annual or semiannual reporting to learners and their families.

**Credentialing**: Awards the Australian Learner Competency Credential (LCC) at key educational transitions, capturing long-term assessments and validating learner competence. This credential is useful beyond the school setting for transitions, admissions, internal qualifications, and employment.

#### 2. Current Challenges

One of our big challenges as the stakes become higher and we need to warrant the credentials we produce is to check if teachers understand our items in the same way for very different contexts. We don't know what evidence teachers use to make their judgement about the competency level of each student and if it is very different from one subject to another even though the questions are about a general capability e.g., collaboration. Our current challenge lies in the nuanced understanding and assessment of complex competencies such as collaboration, communication, and ethical decision-making. These competencies are essential for students to thrive beyond the classroom, yet their subjective nature poses a significant challenge in creating a standardized assessment framework.

#### 3. Problem

How to ensure that teachers understand each topic in the same way?

How to know what evidence the teacher uses to judge each student's ability?

How to check the difference of ability level between different disciplines？

#### 4. Potential solutions

Potential solution 1: the function of "comprehensive evaluation of students"-this function integrates the scores of teachers in all subjects and calculates the average score of each student in different abilities (such as communication ability and learning ability). By averaging the scores of teachers' questionnaires in various subjects, the consistency of teachers' understanding of problems and evaluation criteria is solved to some extent, and the authority and accuracy of evaluation are enhanced. At the same time, this function also provides us with the possibility of further analysis and evaluation basis and improvement of evaluation system. Potential solution 2: manual weight assignment-before making the questionnaire, ask the teacher to assign weights to each evaluation standard (such as communication accounting for 10%, personal and social capability accounting for 20%, etc.). So as to standardize the understanding of each topic and ensure that each teacher has made clear the relative importance of each evaluation standard when evaluating. This can not only improve the consistency of teachers' evaluation, but also reflect students' performance on different topics and abilities more accurately.