



平成30年8月7日
教育課程部会
児童生徒の学習評価に関する
ワーキンググループ
資料2

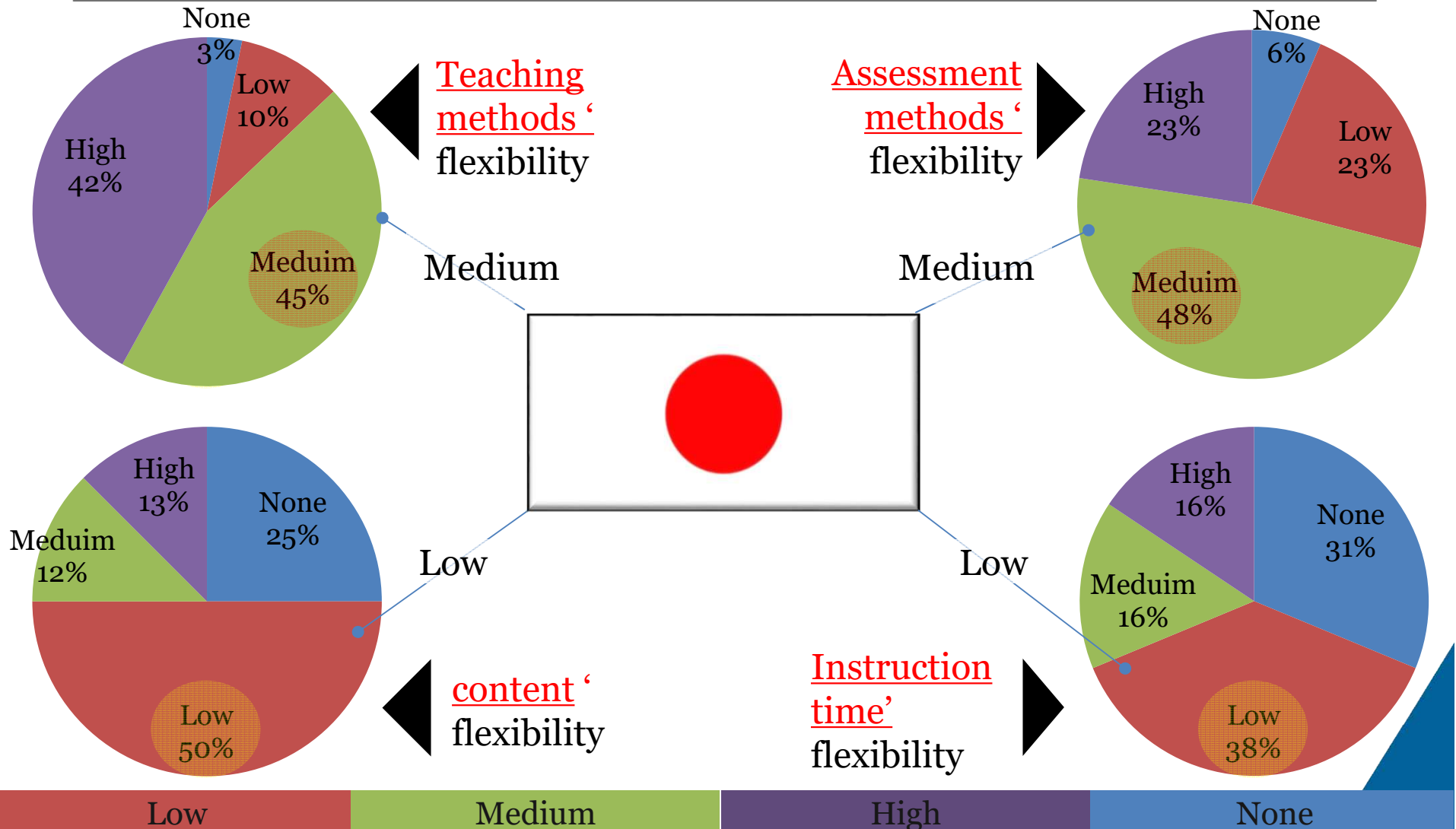
Preliminary findings from the OECD Future of Education and Skills 2030 Construct Analysis: Assessment of “attitudes” and “social & emotional skills”

Miho Taguma
Senior Policy Analyst/ project manager of OECD
Education 2030
Directorate for Education and Skills
OECD



Background: Teachers/schools' flexibility (work in progress)

- **Low flexibility** on content and instruction time, but **high flexibility** on teaching and assessment methods



ARG, AUS, BRA, CAN, CHL, CHN, CRI, CZE, DNK, EST, FIN, HKG, HUN, IND, IRL, JPN, KAZ, KOR, MEX, NZL, IRL, (UK), NOR, POL, PRT, RUS, SCT (UK), SGP, ZAF, NLD, TUR, USA. Updated to the beginning of 2018.



Q1. How teachers/ schools are using the flexibility in different countries?

Selected examples of attitudes/ social & emotional skills:

- ***Curiosity***
- ***Learning to learn / meta-learning***
- ***Motivation***
- ***Collaboration***
- ***Respect***





General observations

- Of the 5 constructs:
 - Those that are **personal and internal** (*curiosity, motivation, learning to learn*) are typically assessed by validated self-report.
 - Those that have **inter-personal traits** (*collaboration, respect*) are assessed through various means, e.g. behaviour observations, peer-evaluations, performance-based test.





Curiosity (work in progress)

Definition: Pre-disposition to recognize and search for new knowledge and experiences (OECD, 2016).

- Assessment instruments: Examples

Instrument	Type	What does it measure?	References
Melbourne Curiosity Inventory (AUS)	Self-report	Assessment measuring trait-curiosity and state-curiosity. E.g. "New situations capture my attention"; "I am curious about things" Responses are given on a 4-point scale from "Almost never" to "Almost always"	Naylor, 1981
National Education Monitoring Project (NZL)	Self-report for students' attitudes (performance-based tasks were administered for subjects' knowledge and skills for national monitoring but Teachers and principals have access to NEMP task materials and marking instructions, so that they can use them within their own schools	NEMP conducted annual assessments of student achievement, values, and attitudes at Year 4 and Year 8; questionnaire investigated students' interests in and attitudes to selected subjects. Curiosity is one of the attitudes included in various subjects. Assessment tasks are performance-based, ranging from one-to-one interviews, work stations and teamwork that will require students to transfer learning to authentic close-to-real life situations. NEMP was designed not as high stakes for students; it aimed to make tasks meaningful and enjoyable to them. The primary purpose of this assessment is to collect rich information on the processes students use to solve problems or conduct experiments.	http://nemp.otago.ac.nz/index.htm
Curiosity and Exploration Inventory	Self-report	A 7-item scale designed to measure respondents' recognition, pursuit, and integration of new and challenging stimuli and experiences. The scale is divided into two factors: (1) Exploration (pursuing novelty) and (2) Absorption (being absorbed in activities). Respondents answer each item using a 7-point Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree).	Kashdan, Rose, & Fincham, 2004;



Learning to learn / Meta-learning (work in progress)

Definition: The state of "being aware of and taking control of one's own learning": meta-learning (John Biggs, 1985). The ability and willingness to adapt to novel tasks, activating one's commitment to thinking and the perspective of hope by means of maintaining one's cognitive and affective self-regulation in and of learning action.

- Assessment instruments: Examples

Instrument	Type	What does it measure?	References
European test learning to learn framework (Europe – FRA, ITL, SLN, AUT, ESP, POR, CYP)	Self-report	Measures "learning to learn" skills: Cognition, Metacognition and affective dimensions	(European Council, 2000)
CCST (NLD)	Self-report	Measures cross-curricular skills and "learning to learn" skills	Elshout-Mohr et.al., 2004
Effective Lifelong Learning Inventory (ELLI) (Europe)	Self-report	Measures a person's orientation towards effective lifelong learning	Deakin, Broadfoot and Claxton. 2004.
"Life as Learning (LEARN)" (International, SWE, FIN)	Self-report	Measures "learning to learn" skills	Hautamäki et al., 2002
Learning to Learn National Assessment (FIN)	National assessments sample (self-report)	Measures to evaluate "learning to learn" skills" for Years 3, 6 and 9 of compulsory education.	the Centre of Educational Assessment at the University of Helsinki



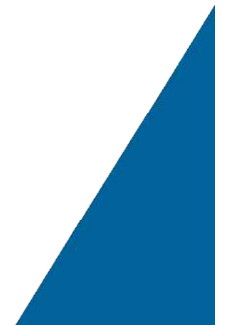


Motivation (work in progress)

Definition: Internal process, state, or force that moves an individual to act (Wigfield et al., 2015).

- Assessment instruments: Examples

Instrument	Type	What does it measure?	References
Academic Self-Regulation Questionnaire (International)	self-report	measure assessing students' autonomous vs. controlled motivation in academic contexts . E.g. "Why are you studying in general? I'm studying because I'm supposed to." Responses are given on a 5-point scale from "Not important" to "Very important"	Ryan & Connell, 1989
The Science Motivation Questionnaire (Self-report	Assess the motivation of intermediate school students in Islamabad towards science through a 30-item questionnaire, ranging from 0 (strongly agree) to 5 (strongly disagree). Ex: My personal goals and objectives associate with my science learning.	<u>Mubeen and Raid, 2014</u>





Collaboration (work in progress)

Definition: Co-ordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem” (Roschelle and Teasley, 1995, p. 70 in PISA CPS Framework, 2017)

- Assessment instruments: Examples

Instrument	Type	What does it measure?	References
Behavioral observation	behavioural observation scale	Focused on interpersonal skills and self-management skills. derived from critical incidents to provide context relevant examples. Each member of the team rated each other team member on items related to the following 13 different dimensions: 1. Reaction to conflict, 2. Addresses conflict, 3. Averts conflict, 4. Synthesis of team’s ideas, 5. Involving others, 6. Effective communication, 7. Goal-setting/achievement, 8. Team citizenship, 9. Commitment to team, 10. Focus on task-at-hand, 11. Preparation for meetings, 12. Providing/reaction to feedback, 13. Performance management	Taggar and Brown (2001)
The Comprehensive Assessment of Team Member Effectiveness instrument	peer evaluation	With this instrument, peers anonymously rate each other based upon their experience in the team interaction. relies upon Likert-type scales for rating team members on questions relating to four dimensions: Contributing to the team’s work Interacting with teammates Keeping the team on track Expecting quality Having relevant knowledge, skills and abilities	Loughry, Ohland and Moore, 2007; Taggar and Brown, 2001
PISA (International)	Cognitive testing	Matrix for collaboration includes: (1) Establishing and maintaining shared understanding, (2) Taking appropriate action to solve the problem, and (3) Establishing and maintaining team organisation. Collaborative skills include: Grounding, Explanation, Co-ordination, Filling roles, Perspective taking, Audience design, Argumentation, Mutual regulation	OECD, 2017. https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf



Respect (for self, others including cultural diversity) (work in progress)

Definition: Attitude towards someone or something (e.g., a person, a belief, a symbol, a principle, a practice, etc.) where the object of that attitude is judged to have some kind of importance, worth or value which warrants positive regard and esteem (the Council of Europe)

- Assessment instruments:
 - Very few self-report instruments are targeted for children and teenagers. When available, they are used to measure respect although those are in specific contexts such as respect in close relationships, respect at work, and self-respect.
 - Respect for others is best measured through others' perspectives instead of relying on self-report instruments, such as through observational tools (such as utilising rubrics) which document behaviour and communication, as well as through reflection assignments using appropriate prompts.





***Q2. How do countries support teachers/
schools to use the flexibility in the most
meaningful ways for students' enhanced
learning?***

Some considerations....





1. Clarifying **the purpose of assessment** considering the nature of “attitudes/ social & emotional skills”.

Assessment for Learning
(formative assessment)
- Low stake

Cf.

Assessment of Learning
(summative assessment)
- High stake

- Classroom-based formative assessment is often defined as **a frequent, interactive assessment of student progress to identify learning needs and shape teaching.**
 - **Frequency** – to support students’ learning by providing relevant and timely feedback (NZL)
 - **Sensitivity on wording:** from "reporting" to "communicating student learning“ (BC/ CAN)
- A growing number of countries have either a **policy framework** to promote formative assessment in classroom at (e.g. Australia, French community of Belgium, Denmark, Estonia and many others), **align it with curriculum framework** (e.g. Finland), or have **formative assessment systems** (e.g. Netherlands).
- For implementation, it is left to schools in most counties while Australia, Korea and Spain also embed it part of **initial teacher education programmes** and Estonia **requires schools to report their implementation strategy** to promote formative assessment.



Country examples: Student formative assessment frameworks

Country	Is there a policy framework for promoting student formative assessment in the classroom?	What requirements are part of the policy framework for promoting student formative assessment in the classroom?
	1	2
Australia	ISCED levels 1, 2 and 3 (general): Yes, at the state level for all schools; ISCED level 3 (pre-voc and voc): No	ISCED level 1: For schools to implement strategies for student formative assessment; ISCED levels 2 and 3: For schools to implement strategies for student formative assessment; for student formative assessment to be part of initial teacher education programmes
Belgium (Fr.)	Yes, at the central level for all schools	For schools to implement strategies for student formative assessment
Canada	Varies across provinces/territories	Varies across provinces/territories
Denmark	Yes, at the central level for all schools	For schools to implement strategies for student formative assessment ²
Estonia	Yes, at the central level for all schools	For schools to implement strategies for student formative assessment; for schools to report on their strategies to promote student formative assessment
Iceland	ISCED levels 1 and 2: yes, at the central level for all schools ⁴ ; ISCED level 3: no	For schools to implement strategies for student formative assessment
Ireland	ISCED level 1: yes, at the central level for public schools and government-dependent private schools only; ISCED levels 2 and 3: no ⁵	For schools to implement strategies for student formative assessment
Israel	ISCED levels 1 and 2: yes, at the central level for all schools; ISCED level 3: no	For schools to implement strategies for student formative assessment



Italy	Yes, at the central level for all schools ⁶	For schools to implement strategies and criteria for student formative assessment
Korea	Yes, at the central level for all schools	For schools to implement strategies for student formative assessment; for student formative assessment to be part of initial teacher education programmes; for teachers to undertake professional development in student formative assessment
Luxembourg	Yes, at the central level for public schools only	For schools to implement strategies for student formative assessment
Mexico	ISCED 1 and 2: yes, at the central level for all schools; ISCED 3: yes, at the central level for centrally managed public schools, at the state level for locally managed public schools ⁷	For schools to implement student formative assessment; for student formative assessment to be part of initial teacher education programmes ⁸
New Zealand	Yes, at the central level for state schools and state-integrated schools only	For schools to implement strategies for student formative assessment
Norway	Yes, at the central level for all schools ¹⁰	For schools to implement student formative assessment
Poland	ISCED levels 1, 2 (general), ISCED level 3: yes, at the central level for all schools; ISCED level 2 (pre-voc and voc): No	For schools to implement student formative assessment ¹¹
Portugal	ISCED 1, 2 (general) and 3 (general): yes, at the central level for all schools	For schools to implement strategies for student formative assessment (approved by the pedagogical council at each school)
Slovenia	Yes, at the central level for all schools	For schools to implement student formative assessment
Spain	ISCED levels 1, 2 (general) and 3: yes, at the central level for all schools; ISCED level 2 (pre-voc and voc): no	For schools to implement student formative assessment; for schools to report on their strategies to promote student formative assessment; for student formative assessment to be part of initial teacher education programmes; for teachers to undertake professional development in student formative assessment
United Kingdom (Northern Ireland)	Yes, at the central level for all schools	For schools to implement strategies for student formative assessment

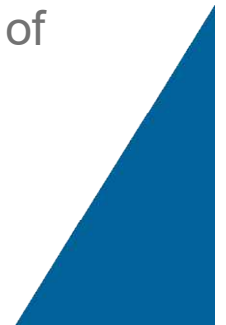


2. Consider the use of **multiple sources of assessments** for (a) “triangulation” – to ensure fairness, validity & reliability, and (b) “student agency” in assessment through engaging students in assessments (self- and peer-assessments)

Teacher-based assessments:

- Validity can be high especially more complex competencies can be assessed through continuous teacher-based assessment.
- Reliability is low (inter-rater & intra-rater reliability)
- Strategies to improve the reliability include using scoring guides, negotiated scoring criteria, external benchmarks, training for teachers, multiple judgements and external moderation.

In the Flemish Community of Belgium, the central education authorities are promoting a shift towards a “broad assessment culture” focusing on **formative assessment and new assessment approaches, including observation, portfolios, reflection sheets and self- and peer-assessment**. It is also being highlighted that it is more important to report on student progress over a certain time period rather than on absolute performance (Flemish Ministry of Education and Training, 2010).





In Slovenia, the principles for assessment and examinations are specified in Rules on Examination and Assessment of Knowledge for the different levels of education. Among other things these principles require teachers :

- **to use a variety of forms and methods of verification** and assessment of knowledge;
 - to allow students **a critical reflection and insight into the acquired knowledge** and to contribute to the **democratisation of relations between students and teachers** (Brejc et al., 2011).
 - Portugal align curriculum reform with **assessment focused on motivating students, giving high quality feedback, and including the active participation of learners in the assessment process**. The use of approaches beyond written tests, such as a performance assessment, puts the learner and learning at the centre of the assessment process. Some schools use **peer assessments not only individual peers but also collective work**.
- British Columbia Canada acknowledges the importance of **self-assessment** and flexibility in the ways that **students can demonstrate their learning (e.g., documentation, samples of work, portfolios)**.



Examples of using multiple sources to assess social and emotional skills (OECD Study on Social and Emotional Skills)

Co-operation	Persistence	Optimism	Creativity
Trust	Responsibility	Stress Resistance	Self-efficacy
Empathy	Sociability	Emotional Stability	Critical Thinking
Self-control	Assertiveness	Curiosity	Metacognition
Achievement Motivation	Energy	Tolerance	

STRUCTURE OF INSTRUMENTS		Students	Parents	Teachers	Principals
Assessment of students' skills	Indirect assessment	✗	✓	✓	✗
	Direct assessment	✓	✗	✗	✗





3. Design support for teachers and schools

- British Columbia, Canada provide guidelines **for teachers to support student self-assessment**
- In the Flemish Community of Belgium, several specific evaluation instruments have been developed by various institutions **to help teachers assess non-cognitive performances of their students**. One of the most commonly used tools is the SAM-scale (Scale for Attitude Measurement), which has been developed as a tool to support teacher to determine to what extent a pupil scores high or low for certain attitudes, e.g. flexibility, diligence and responsibility.
- New Zealand's National Education Monitoring Project **allow some performance-based tasks, including students' attitudes, to be used by teachers**.





4. Consider closing the **time lag** between curriculum redesign (competency-based) and assessment practices (still traditional focus on knowledge with some skills)

- Competency-based assessments, such as performance-based assessments, are likely to be more advanced in the vocational education and training (**VET**) sector or in hiring practices in the **private sector/companies** and holistic assessments are more widely practised in the early childhood education and care (**ECEC**) sector.

In Finland, for example, all vocational qualifications include skills demonstrations, which form part of student assessment.

- Future consideration for use of ICT - (1) the “migratory” strategy (i.e. using ICT to deliver traditional assessment formats more effectively and efficiently) and (2) the “transformative” strategy (i.e. using ICT to change the way competencies are assessed) e.g. simulation, interactivity and student’s actions stored in logfiles to used to track students’ inquiry trajectories. This could offer new avenues for assessing broader competencies on a larger scale, with possibly more affordable costs. But further studies on the validity and reliability of different automated essay scoring tools are much needed.

