

# Teaching and learning informed by assessment in the Diploma Programme

Guide and teacher support material

# Teaching and learning informed by assessment in the Diploma Programme

Guide and teacher support material

# Diploma Programme

## Teaching and learning informed by assessment in the Diploma Programme

Published June 2021

Published on behalf of the International Baccalaureate Organization, a not-for-profit educational foundation of 15 Route des Morillons, 1218 Le Grand-Saconnex, Geneva, Switzerland by the

International Baccalaureate Organization (UK) Ltd  
Peterson House, Malthouse Avenue, Cardiff Gate  
Cardiff, Wales CF23 8GL  
United Kingdom  
Website: [ibo.org](https://ibo.org)

© International Baccalaureate Organization 2021

The International Baccalaureate Organization (known as the IB) offers four high-quality and challenging educational programmes for a worldwide community of schools, aiming to create a better, more peaceful world. This publication is one of a range of materials produced to support these programmes.

The IB may use a variety of sources in its work and checks information to verify accuracy and authenticity, particularly when using community-based knowledge sources such as Wikipedia. The IB respects the principles of intellectual property and makes strenuous efforts to identify and obtain permission before publication from rights holders of all copyright material used. The IB is grateful for permissions received for material used in this publication and will be pleased to correct any errors or omissions at the earliest opportunity.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the IB's prior written permission, or as expressly permitted by the [Rules for use of IB Intellectual Property](#).

IB merchandise and publications can be purchased through the [IB Store](#) (email: [sales@ibo.org](mailto:sales@ibo.org)). Any commercial use of IB publications (whether fee-covered or commercial) by third parties acting in the IB's ecosystem without a formal relationship with the IB (including but not limited to tutoring organizations, professional development providers, educational publishers and operators of curriculum mapping or teacher resource digital platforms etc) is prohibited and requires a subsequent written license from the IB. License requests should be sent to [copyright@ibo.org](mailto:copyright@ibo.org). More information can be obtained on the [IB public website](#).

## IB mission statement

The International Baccalaureate aims to develop inquiring, knowledgeable and caring young people who help to create a better and more peaceful world through intercultural understanding and respect.

To this end the organization works with schools, governments and international organizations to develop challenging programmes of international education and rigorous assessment.

These programmes encourage students across the world to become active, compassionate and lifelong learners who understand that other people, with their differences, can also be right.





# IB learner profile

The aim of all IB programmes is to develop internationally minded people who, recognizing their common humanity and shared guardianship of the planet, help to create a better and more peaceful world.

As IB learners we strive to be:

## INQUIRERS

We nurture our curiosity, developing skills for inquiry and research. We know how to learn independently and with others. We learn with enthusiasm and sustain our love of learning throughout life.

## KNOWLEDGEABLE

We develop and use conceptual understanding, exploring knowledge across a range of disciplines. We engage with issues and ideas that have local and global significance.

## THINKERS

We use critical and creative thinking skills to analyse and take responsible action on complex problems. We exercise initiative in making reasoned, ethical decisions.

## COMMUNICATORS

We express ourselves confidently and creatively in more than one language and in many ways. We collaborate effectively, listening carefully to the perspectives of other individuals and groups.

## PRINCIPLED

We act with integrity and honesty, with a strong sense of fairness and justice, and with respect for the dignity and rights of people everywhere. We take responsibility for our actions and their consequences.

## OPEN-MINDED

We critically appreciate our own cultures and personal histories, as well as the values and traditions of others. We seek and evaluate a range of points of view, and we are willing to grow from the experience.

## CARING

We show empathy, compassion and respect. We have a commitment to service, and we act to make a positive difference in the lives of others and in the world around us.

## RISK-TAKERS

We approach uncertainty with forethought and determination; we work independently and cooperatively to explore new ideas and innovative strategies. We are resourceful and resilient in the face of challenges and change.

## BALANCED

We understand the importance of balancing different aspects of our lives—intellectual, physical, and emotional—to achieve well-being for ourselves and others. We recognize our interdependence with other people and with the world in which we live.

## REFLECTIVE

We thoughtfully consider the world and our own ideas and experience. We work to understand our strengths and weaknesses in order to support our learning and personal development.

The IB learner profile represents 10 attributes valued by IB World Schools. We believe these attributes, and others like them, can help individuals and groups become responsible members of local, national and global communities.

# Contents

<b>Executive summary</b>	<b>1</b>
Executive summary	1
<b>Introduction: What is teaching and learning informed by assessment and why is it important?</b>	<b>4</b>
The DP's approach to assessment	4
Teaching and learning informed by assessment: Summary of the literature	6
<b>What makes feedback effective and how do we deliver meaningful feedback?</b>	<b>10</b>
What makes feedback effective?	10
Developing student evaluative judgment through feedback	12
Should high-stakes summative assessment criteria be used for formative feedback?	14
Examples of effective practices from DP educators	15
<b>How do we design meaningful assessments that support formative feedback?</b>	<b>17</b>
Introduction	17
Universally-designed assessments: Varied and multimodal	18
Feedback that reflects current achievement, learning progress or process	19
Aspects of validity in formative assessment and feedback	22
<b>Common types of formative assessment in DP classrooms and what they aim to measure</b>	<b>27</b>
Three main types of effective designs in DP classrooms	27
Is "teaching to the test" quality formative assessment?	31
<b>How do we assess conceptual understanding in IB classrooms?</b>	<b>32</b>
Inquiry-based and concept-based curriculum	32
What are concepts and how do we assess conceptual understanding?	33
<b>What do classroom assessments look like in a skills-based curriculum?</b>	<b>43</b>
The development of skills within a concept-based curriculum	43
An example of a skills-based approach in DP business management	46
<b>How do we make assessments fair and inclusive for all learners?</b>	<b>51</b>
Universally-designed assessments	51

<b>How do we support IB learners in becoming assessment capable?</b>	<b>55</b>
Student agency in assessments: Peer- and self-assessments	55
Metacognition and self-regulated learning in formative assessments	58
Student-generated questions and assessments	60
Promising practices in IB schools: Examples of student-led assessment	61
<b>Using assessments to support student development of ATL skills and learner profile attributes</b>	<b>62</b>
Formative assessments and the ATL skills	62
Formative assessments and the IB learner profile	65
How can digital technology be used to enhance and transform formative assessments?	67
<b>How can a positive learning environment that promotes formative assessment be cultivated?</b>	<b>71</b>
Supporting IB teachers to become assessment-capable	71
School policies and practices that support quality formative assessments	72
<b>Appendices</b>	<b>74</b>
Glossary	74
References	79

## Executive summary

International Baccalaureate (IB) World Schools, educators and students bring a wealth of knowledge and rich and diverse assessment practices to the IB community. The International Baccalaureate Organization (IBO)'s Assessment Division has published a series of policies and documents that map out a comprehensive and diverse set of principles and practices on valid assessments in the IB. To date, many of the examples published have focused on summative assessments designed and assessed or moderated by the IB. To build on the collective wisdom and community of practices in IB World Schools on classroom assessments, a pilot study on "Effective practices in formative assessment in the DP" (Diploma Programme) was conducted from October 2018 to June 2020. The aim of this study is to understand common aspirations and interests in formative assessments, to explore key questions that emerge, and to gather and build on examples of effective practices from IB schools across regions, school and cultural contexts.

This combined guide and teacher support material is based on findings and examples from the pilot study, and integrates them with recent literature. The aim is to map out a set of principles and practices that are driven by research and embody the IB's education philosophy. One of the key features of formative assessment in IB classrooms is that it is driven by teacher and student agency and its responsiveness and dialogic nature. Therefore, this publication is not meant to be a prescriptive guide. The principles outlined here could underpin effective practices of assessment **for** and **as** learning in IB classrooms, while always respecting and attending to the agency and self-leadership demonstrated by IB educators and learners in making decisions about assessment that are best suited to their learning and teaching contexts.

Being mindful of the diverse and intercultural nature of the IB's international education, the goal is to present a range of practices and examples on formative assessment from IB schools in different contexts. Nevertheless, what has been captured here may be a small part of what is possible. The aim, therefore, is that this publication will help to initiate a meaningful dialogue in the IB community, and create a common language to discuss and construct the way forward, to celebrate and promote quality assessments for formative purposes in IB classrooms.

The themes and questions that emerged in the pilot study and in ongoing conversations with IB educators are organized into 11 key questions in this document. These are the most common big questions that are found when exploring effective principles and practices on formative assessments in the IB. The guide and teacher support material is organized around these 11 key questions, with examples of promising practices integrated throughout.

Note that the publication is not designed to be read from beginning to end. Teachers and coordinators can read the sections they feel are most relevant to their needs, or follow it in any order.

## Eleven questions on teaching and learning informed by assessment in the Diploma Programme

1. What is teaching and learning informed by assessment and why is it important?
2. What makes feedback effective and how do we deliver meaningful feedback?
3. How do we design meaningful assessments that support formative feedback?
4. What are common types of formative assessments in Diploma Programme (DP) classrooms and what do they aim to measure?
5. How do we assess conceptual understanding in IB classrooms?
6. What do classroom assessments look like in a skills-based curriculum?

7. How do we make assessments fair and inclusive for all learners?
8. How do we support IB learners in becoming assessment capable?
9. How do we use assessments to support student development of approaches to learning skills (ATL) and the learner profile attributes?
10. How can digital technology be used to enhance and transform formative assessments?
11. How can a positive learning environment that promotes formative assessment be cultivated?

## Methodology and limitations

The insights and examples in this publication are drawn from the pilot study on “Effective practices in formative assessment in the DP”, conducted from October 2018 to June 2020. The sources of data and information in the pilot study are based on but not limited to the following.

1. A review of the literature on research and academic work carried out on formative assessments (ongoing).
  - a. Peer-reviewed articles from 2000 to 2020 and earlier seminal works.
  - b. Books, journals and other related articles.
  - c. Key search terms: “formative assessments”, “assessment for learning”, “assessment as learning” and other related concepts and strategies.
2. Case studies of six IB World Schools (school visits conducted in October 2018 to January 2020).
  - a. Class visits across a range of DP subjects and the core.
  - b. Focus-group discussions and individual interviews with DP teachers, coordinators and school leaders.
3. Online discussions on Basecamp (ongoing).
4. Interviews and focus-group discussions conducted with IB members of staff (May to October 2018).
5. Internal research report on “A dialogue on formative assessment in the International Baccalaureate Diploma Programme” (shared internally with IB members of staff in July 2019).

One of the limitations of the pilot study is that all the examples presented focus only on the DP. While the general principles and approaches discussed closely align with the IB’s education philosophy and could be relevant to formative assessment practices across the IB continuum, the examples presented may not be able to address fully the unique programme features of the Primary Years Programme (PYP) and Middle Years Programme (MYP).

Moreover, the current pilot study is limited in its scope. The IB World Schools and educators that participated in this pilot study were recruited through IB regional conferences or were recommended as schools that demonstrate effective practices on formative assessments or that have focused on formative assessment as a key area of interest in their school’s development. While the number of schools and educators that participated was small, it is also the case that schools from a diverse range of contexts were deliberately sought out, including the following.

- State schools, private national schools and private international schools
- Schools that mainly used English as the language of instruction, and a bilingual school that used both French and English as the languages of instruction
- IB continuum schools where many teachers taught across the MYP and DP. Similarly, schools that only offered the DP or combined the IB programmes with other national or local curriculums in earlier years or grades (for example MYP 1–3 years)
- Schools that only offered the DP in high school, and those that offered both the DP and the state or local curriculum in the high school or diploma level
- Schools across three geographical regions, namely, Asia, Europe and the Americas

However, the number of case studies and the scope of the pilot study are limited. Other limitations of this document may also include unintended researcher bias.

There is, therefore, no suggestion that this publication encompasses a full and representative sample of the rich and diverse practices of effective formative assessments already found, or that may be found in future, in IB World Schools. There are many IB schools and educators who were not able to take part in this study, but are enthusiastic about formative assessment and doubtless would like to share their promising practices. The IB encourages educators to challenge any of the ideas discussed here, and to share with us and the IB community effective practices from their contexts.

### **Acknowledgments**

The IB would like to express its gratitude to all of the IB educators who engaged in discussions that contributed to this publication, sharing many fascinating examples of effective practices. A special thank you goes to the DP students, teachers and school leaders in the following schools, where class visits took place that inspired many of the ideas in this publication, based on the learning and teaching practices found there.

- Branksome Hall, Canada
- Hockerill Anglo-European College, United Kingdom
- King Edward's Witley, United Kingdom
- Międzynarodowe Szkoły Paderewski w Lublinie, Poland
- NIST International School, Thailand
- Toronto French School —Canada's International School, Canada

## The DP's approach to assessment

Assessment plays a crucial role in promoting learning and growth in students and in eliciting evidence of learning and tracking progress. The IB's approach to assessment is that learning, teaching and assessment should effectively inform and support one another. The IB's vision is that students will benefit from a holistic assessment experience throughout their IB education, and that their sense of self-efficacy and agency are enhanced in the process. While the four IB programmes (PYP, MYP, DP and Career-related Programme) have different emphases and balances in assessment in support of their curricular goals, and are developmentally appropriate, the underlying principles of assessment are essentially the same across programmes.

In the IB, quality assessments are seen as:

- relevant—supporting learning objectives and curriculum outcomes
- meaningful—assessing what is important and not just what is easy to measure
- authentic—reflecting ways in which our students may encounter these activities in the real world
- universally designed:
  - construct relevant—assessing what we intend to assess
  - varied—using a wide range of assessment tasks and activities that will support diverse learners and all types of learning
  - multimodal—encouraging students to express their skills and knowledge in multiple ways, and designing assessment tasks and activities to allow multiple means of representation
  - fair—do not lead to an unintended advantage or disadvantage to any learner
  - inclusive—allowing meaningful engagement of all learners
- stimulating—extending learning and considering wider student competencies and higher-order cognitive skills
- valid—balancing important principles, such as construct relevance, fairness, manageability, reliability and comparability, to reach the intended purpose of assessment
- fit for purpose—designed, implemented and interpreted in a consistent way to fulfil the purposes for which they are intended (assessments with different purposes may be designed, analysed and interpreted in different ways)
- aligned with IB education philosophy—supporting the IB education philosophy and in line with the ATL and the learner profile.

As can be seen from the title of this document, *Teaching and learning informed by assessment in the Diploma Programme*, the principles and effective practices discussed throughout sit neatly with the IB approaches to teaching and learning. This is because discussions regarding classroom assessments and formative feedback need to go hand in hand with the pedagogical approaches and philosophy of an IB education. For the same reason, while the examples in this publication focus on the IB DP as the highest stakes programme in the IB (and where perhaps there is the narrowest view of formative assessment), the content of this publication may also be relevant to the MYP and PYP.

For further details on the IB's approach to assessment and how to balance these assessment and pedagogical principles in practices, please refer to the following IB publications.

- [\*Assessment principles and practices—Quality assessments in a digital age\*](#)
- [\*What is an IB education?\*](#)
- [\*Programme standards and practices\*](#)
- [\*Diploma Programme: From principles into practice\*](#)

- *Diploma Programme: Approaches to teaching and learning website*

In addition to these key resources, the IB provides a range of other resources that give focused guidance to teachers, coordinators and other stakeholders, based on the principles set out in these key publications. These can be found on the IB website and the programme resource centre.



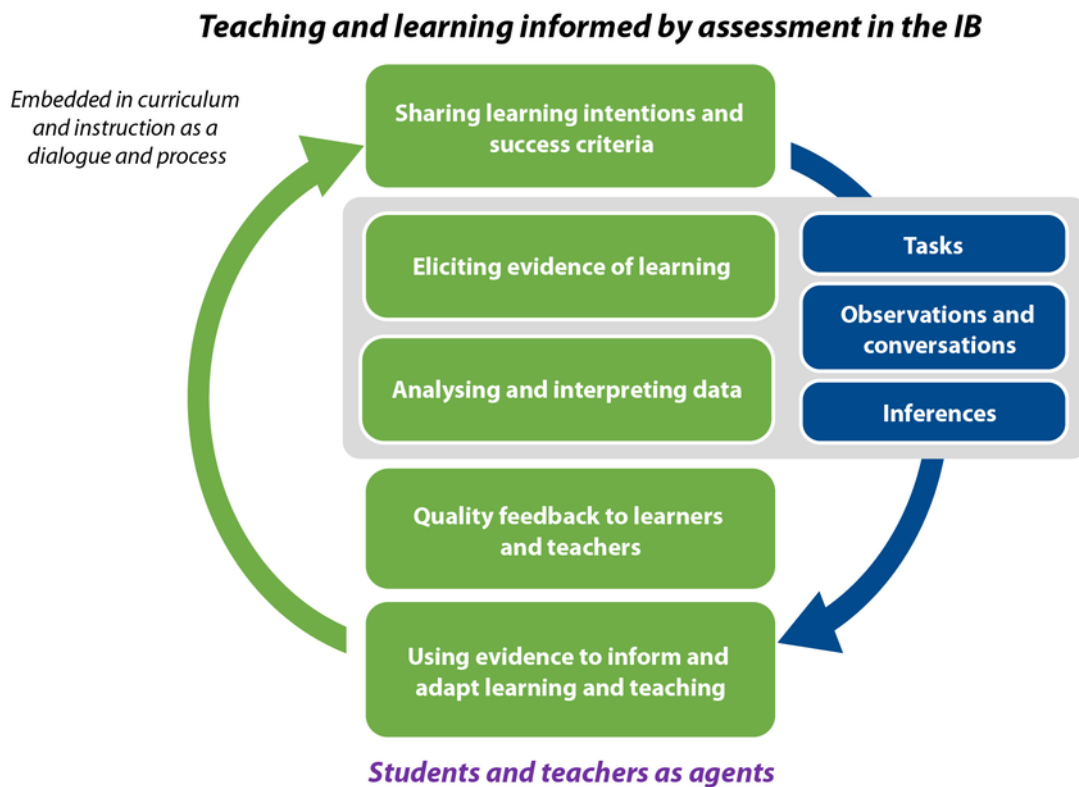
## Teaching and learning informed by assessment: Summary of the literature

This section outlines a conceptual framework on teaching and learning informed by assessment in the IB by summarizing some of the key literature. Each topic will be explored further in the different sections of this publication.

As seen in Figure 1, formative assessment is assessment that is embedded in curriculum and instruction, where the inferences drawn from assessment outcomes “relate to the kinds of actions that would best help the student learn” (Black, Wiliam 2018: 553).

Figure 1

*A model for teaching and learning informed by assessment*



The IB keeps abreast of the latest research and global best practices in formative assessment, sometimes known as assessment **for** or **as** learning. The following are some key points from IB research and important peer-reviewed articles.

For formative assessments to be meaningful and effective, the IB recognizes that:

- assessments are tasks, conversations, observations and inferences
- quality assessments need to be embedded in curriculum and instruction
- formative assessments are founded on research in pedagogy and learning sciences
- the key is to develop agency and assessment-capable IB learners and teachers.

## Assessment as tasks, conversations, observations and inferences

- “Formative Assessment is Assessment” (Bennett 2011: 20), and assessment is about making inferences on student learning.

### Differences between formative and summative assessments

- The distinction between formative and summative assessments is in the purposes of assessment and “the kinds of inferences being drawn from assessment outcomes” (Black, Wiliam 2018: 553).

	Purpose	Inference
<b>Summative assessments</b>	<ul style="list-style-type: none"> <li>• To demonstrate the competency or level of achievement of a student generally at the end of a course of study or a unit of learning</li> <li>• To determine the readiness of the student to progress to the next stage of education</li> </ul>	<ul style="list-style-type: none"> <li>• Focus on what the student can do at the end of a course of study or a unit of learning.</li> </ul>
<b>Formative assessments</b>	<ul style="list-style-type: none"> <li>• Ongoing assessment during the course of study to provide evidence and feedback that could move learning forward</li> <li>• The primary purpose is to promote student learning in a well-informed meaningful way</li> </ul>	<ul style="list-style-type: none"> <li>• Relate to the kinds of actions that can be taken next which “would best help the student learn” (Black, Wiliam 2018: 553)</li> </ul>

- Formative inferences are those that support learners and teachers to identify strengths and gaps in understanding, such as possible misconceptions and **what steps need to be taken next to move learning forward**.
- Formative assessments consist of three indispensable parts (figure 2).
  - The assessment tasks (can be an activity or a question)
  - The observations and dialogues that take place during the formative assessment process
  - The inferences that can be drawn from the assessment outcomes

Figure 2

Three key elements of formative assessment



- What is observed during formative assessment is not simply students’ answers or the learning product but also the learning process, such as learning strategies and the steps taken to generate those responses.
- It is essential to balance relevant assessment principles in the design, interpretation and evaluation of quality formative assessments to build meaningful and fair assessments.

## Embedding assessment in curriculum and instruction

- Effective formative assessment needs to be embedded in curriculum and instruction, such that it provides insights about learning as it occurs (Shepard 2017).
- Meaningful formative assessment uses the five key strategies shown in the table below.

	Where the learner is going	Where the learner is right now	How to get there
<b>Teacher</b>	Clarifying and understanding learning intentions and criteria for success	Eliciting evidence of student understanding through effective class conversations, observations and learning tasks	Providing feedback to both learners and teachers to move learning forward
<b>Peer</b>		Peer-assessment and feedback: Activating students as instructional resources for one another	
<b>Learner</b>		Student agency: Activating students as owners of their own learning	

Adapted from Wiliam and Thompson (2017)

## Founding assessment on research in pedagogy and learning sciences

- A theoretical framework of formative assessment needs to be founded on research in pedagogy and learning sciences.

**Student agency**—Effective formative assessment promotes students' metacognition and self-regulated learning (Clark 2012; Panadero et al 2018).

**Collaboration**—Formative assessment is collaborative in nature and promotes peer learning and feedback and collaboration with teachers (Strijbos, Wichmann 2018; Topping 2009).

**Feedback**—Effective inferences and feedback are those that are actionable, timely, concrete, forward looking, attend to affect, promote agency, relevant and in some way personalized (Dawson et al 2019; Gibbs, Simpson 2004; Hattie, Timperley 2007; Nicol, Macfarlane-Dick 2006; Yeager, Dweck 2012).

**Conceptual understanding**—Effective formative assessment elicits evidence of conceptual understanding in students and their abilities to apply these skills in novel and unfamiliar contexts.

**Universal design for learning and of assessment**—Quality formative assessments are those that ensure all learners can access and participate in meaningful and challenging learning opportunities. For example, teachers design the assessment task and stimulus such that the information is represented in multiple means. Students are also encouraged to engage with the learning and assessment materials and they are invited to express their knowledge in multiple ways (Hall et al 2012). For further details on universal design of assessments, please refer to the website of the National Center on Educational Outcomes ([https://nceo.info/Assessments/universal\\_design/overview](https://nceo.info/Assessments/universal_design/overview)).

## Agency: Developing assessment-capable IB learners and teachers

- The assessment literacy or assessment capability of educators is key to the effectiveness of assessment for formative purposes (DeLuca et al 2019; DeLuca, Klinger 2010; Xu, Brown 2016).
- In this publication, assessment capability is broadly defined as teachers' "capacity to negotiate assessment knowledge and skills in the context of their classroom teaching" (DeLuca et al 2019: 159).
- Professional development plays an indispensable role in supporting IB teachers to become assessment capable. The development of professional learning communities also provides a space for IB teachers to share and build on professional insights in assessments in IB classrooms.
- There is a need to build in time and space for teachers to share, plan and improve assessment practices in curriculum planning and standardization meetings.
- For IB students to benefit from peer-assessment, self-assessment and feedback, they need to be supported in internalizing success criteria and enhancing their assessment capabilities or "evaluative judgment" (Carless, Boud 2018; Panadero et al 2019). In other words, students are better prepared to understand "what is good" and to reflect on the quality of their work and that of their peers.

## What makes feedback effective?

Effective assessments are those that are fit for purpose, and one of the main purposes of formative assessment is to generate helpful feedback to both learners and teachers.

One question often raised by educators is “what makes feedback effective?”. In general, effective inferences and feedback are those that are actionable, timely, concrete, forward looking, attending to the affective aspect of assessment, promoting student agency, relevant, and in some way personalized (Dawson et al 2019; Gibbs, Simpson 2004; Hattie, Timperley 2007; Nicol, Macfarlane-Dick 2006; Yeager, Dweck 2012).

The following gives some further insights from research on formative feedback in the classroom.

### Feedback and feedforward

**(Hattie, Timperley 2007; Nicol, Macfarlane-Dick 2006)**

- Effective feedback and feedforward inform teachers and learners about current achievement or “where the learner is now”, but also “where to go next” and “how to get there”.
- Feedforward is timely and future oriented. By supporting learners to evaluate past performance, strengths and gaps in understanding, they can then think about what they need to do to improve future performance on upcoming related tasks.

### Constructive criticism

**(Fong et al 2019)**

- While we may have a natural human tendency to prefer positive feedback over negative feedback, receiving negative feedback does not necessarily harm our intrinsic motivation for tasks.
- Constructive criticism can have a positive impact on intrinsic motivation when it provides concrete information and directions for improvement, when it is criterion-related instead of normative (therefore not comparing students to one another), and when it is delivered in person.

### Feedback as dialogic interactions

**(Adie et al 2018; Carless 2015)**

- Effective feedback is often in the format of a dialogue between teacher and student or among peers.
- Effective feedback requires active listening by both teachers and students.
- Teachers provide feedback in the form of questions that ask students to reason, justify, analyse and evaluate their learning.
- Students demonstrate agency by analysing their performance and the feedback, and they respond in terms of the strategies they think will work best for them.
- Students are more willing to articulate their strengths and weaknesses when the teacher seeks their feedback and uses the feedback session to obtain information on the students’ understanding, interpretation and valuing of the feedback.
- The eventual goal is to enhance student ability to self-assess and to promote students’ self-regulated learning.

### Student and teacher perceptions

**(Dawson et al 2019)**

- The purpose of formative feedback is to support improvement instead of justifying a grade.
- Students find feedback more effective when it is usable, sufficiently detailed, relevant to the students' own work, and attendant to the affective aspect of the feedback (for example the impact on motivation or students' self-perception). Teachers tend to focus more on the design of the feedback, such as giving timely feedback, delivered face to face or in the form of a rubric, and giving iterative feedback, as in the case of staged assignments.

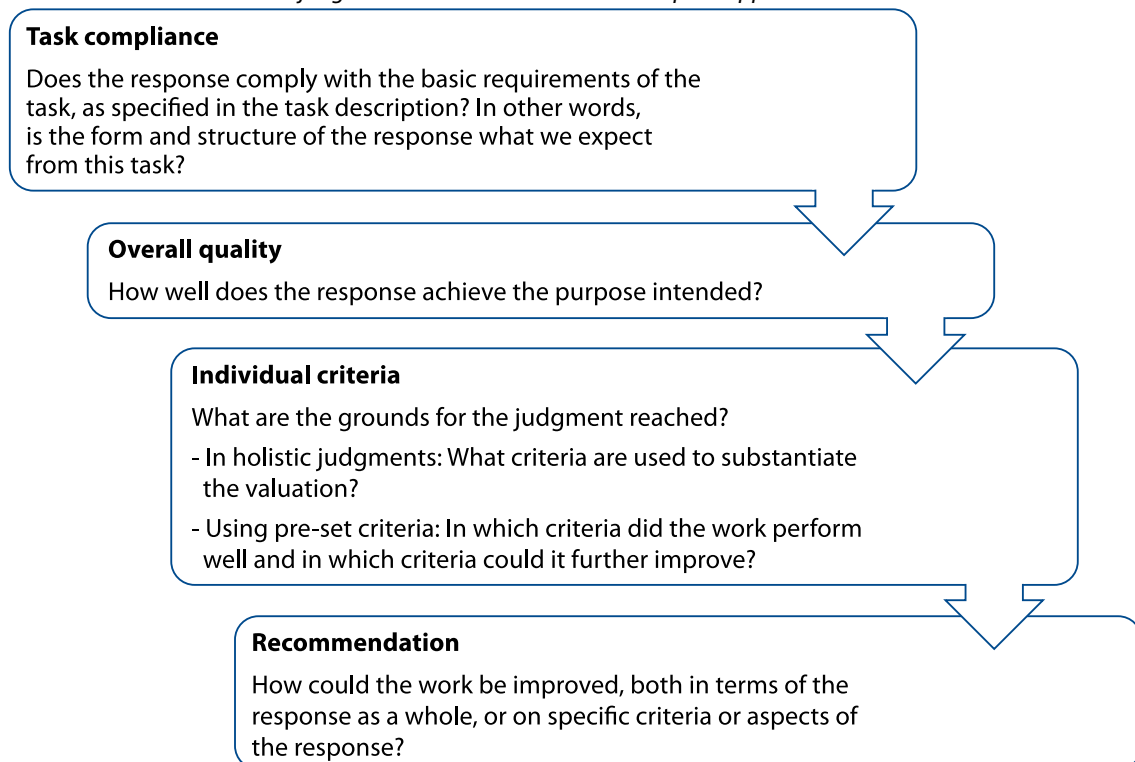
## Developing student evaluative judgment through feedback

- To promote student agency through feedback, it is important to move away from an information transmission model of feedback, where the teacher tells the student about the quality of their work (Sadler 2010).
- Using peer-evaluation and feedback, students gradually develop their evaluative judgment through giving and actively processing feedback. The assessment process itself becomes a learning experience —“assessment as learning”.
- Students recognize and understand the reasons for quality and develop their personal capability to make complex judgments about their work.

According to Sadler (2010), there are four essential elements for peer-appraisal and feedback (figure 3).

Figure 3

*Evaluative judgment: Four essential elements of peer-appraisal and feedback*



- These four elements of task compliance, overall quality, individual criteria and recommendations for improvement are also the essence of assessment capability. Experienced teachers and assessors are often skilful in making professional judgments on these four aspects of a student’s work.
- The purpose of feedback is to support learners who may be novice assessors to develop the ability to judge the quality of their own work, and that of their peers, through practices and engaging with peer-assessment and feedback.

## Comparative judgment as a formative resource

Another formative exercise to develop student evaluative judgment is the use of comparative judgment. Comparative judgment is designed as an alternative and reliable grading method based on the principle that people are better at making comparisons between two pieces of student work (whether one was better than the other) rather than making absolute judgments on quality (by putting a mark on each piece of work). This process draws on the “tacit knowledge” in assessors, “the kind of knowledge which is very hard to express in words” (Christodoulou 2017: 187).

In formative assessment, comparative judgment is used as a valuable peer-assessment resource for sharing feedback and to develop students’ tacit knowledge about quality. How quick the assessment process is and the reliability of the rank order generated are not the most important factors here. Instead, it is essential to guide students to articulate their rationales and criteria used to support each judgment in the process. The purpose of the activity is that through “assessment as learning”, students gradually build a common understanding on “what good looks like” and could use the exercise to generate a list of useful exemplars of student work (Christodoulou 2017).



## Should high-stakes summative assessment criteria be used for formative feedback?

- The success criteria developed for a summative assessment is not always fit for purpose for a formative assessment task or activity. Summative assessment criteria are often designed to help generate the most reliable grade or quantitative feedback on a piece of work.
- Effective criteria in formative assessments are those that support students in peer- and self-assessments and help in providing feedback that is actionable, timely, concrete and forward looking.
- Often, effective formative feedback is qualitative rather than quantitative, as in the case of “gradeless assessments”.
- It has also been argued that the use of holistic marking and rubrics in some cases increases the reliability and validity of scores in a summative assessment. However, a more analytical approach has been reported to facilitate the exchange of more timely and efficient formative feedback among students, peers and teachers (Chaim, Cazarini 2017; Sadler 2009). An exception is when students are guided in a step-by-step way to engage in holistic appraisal and feedback; for example, in the case of using comparative judgment for peer-assessment.
- Here are some questions that teachers consider helpful when making the decision on whether the high-stakes summative assessment criteria is useful for a formative assessment task or activity.
  - Are my students approaching the end of the course?
  - Are they able to demonstrate most of the knowledge, skills and abilities indicated in the summative assessment criteria?
  - Do we want to assess all the key constructs and complex skills assessed in the high-stakes summative assessment, or do we want to focus on specific skills and provide fine-tuned feedback on those?
  - How can I adapt the assessment design and success criteria to increase the inclusivity of the task and ensure the feedback we give is fair and helpful to all students? (Refer to the section, “How do we make assessments fair and inclusive for all learners?”)
  - Are there ways that we can adjust the assessment criteria to enhance its power in providing concrete and actionable feedback and feedforward?
  - Are the success criteria written in a student-friendly language?
  - How can we adjust the assessment criteria so that it facilitates student peer- and self-assessment and supports the development of students’ evaluative judgment? (Refer to the section, “How do we support IB learners in becoming assessment capable?”)

## Examples of effective practices from DP educators

### The use of developmental rubrics for formative feedback

#### Assessing science inquiry work

Figure 4 is an example shared by Jenny Lacey, a DP physics teacher. The developmental rubric was designed to assess formatively students' performance on science inquiry work, and also used to provide written feedback to students on the first draft of their DP sciences internal assessment, focusing on the exploration strand. The developmental rubric was created based on the summative assessment criteria and using Bloom's taxonomy to identify the developmental stages in each criterion. Students demonstrate increasingly higher-order or complex cognitive skills from remembering to applying or understanding, analysing and then evaluating, in each area of performance.

As illustrated in figure 4, the yellow areas indicate the "zone of actual development" of the student while the green areas indicate the "zone of proximal development". Using a best-fit model, the student showed evidence of working at the applying or understanding level in general and this is identified as the overall zone of actual development.

For example, in figure 4, the student has demonstrated the highest level of performance in his or her ability to justify the inquiry (2.1.3 **Explains** why topic is worth exploring using clear reasons); while there was room for improvement in his or her ability to establish the scientific context (2.2.1 **States** a relevant area of physics).

Developmental rubrics is a powerful tool to provide formative feedback to students. A common problem identified with feedback in the form of a mark or abstract statements is that it may point out "where the student is now" but does not generate insights on "where to go next" or "how to get there". A developmental rubric is a great way for students to self-assess and provide peer feedback to one another and to identify skills to further strengthen. For example, the student in this example may set personal learning goals on "describing how to use the materials correctly to collect data" and "justifying why the variables need to be controlled" to deepen his or her scientific inquiry skills.

Figure 4

Example developmental rubric to assess student performance in science inquiry

Evaluating		2.1.3 <b>Explains why</b> topic is worth exploring using clear reasons	2.2.4 <b>Explains</b> how research led to the chosen research question		3.2.4 <b>Justifies</b> why the range for the IV is relevant in the context of the RQ	3.3.4 <b>Justifies</b> why these variables need to be controlled to ensure the collected data is reliable	4.1.4 <b>Evaluates</b> the relative significance of safety issues	Students at this level can develop an inquiry based on background research to explore a specific physics question that is worth exploring.  Students plan to collect sufficient, relevant data to address the research question.
Analysing	1.1.3 <b>Describes</b> RQ clearly including IV, DV and one controlled variable		2.2.3 <b>Discusses</b> background reading relevant to the research question	3.1.3 <b>Describes how to use</b> the materials correctly to collect data	3.2.3 <b>Proposes</b> an appropriate number of repetitions in the context of the methodology and the materials available	3.3.3 <b>Describes</b> how to control other variables to minimize their impact	4.1.3 <b>Describes</b> how to reduce the safety risks	Students at this level can design a safe inquiry to explore a specific physics question.  Students have completed background research into the relevant physics concepts. They plan to collect sufficient, relevant data.
Applying/ understanding	1.1.2 <b>Includes</b> variables in the research question or title	2.1.2 <b>Identifies</b> independent and dependent variables relevant to the context of inquiry	2.2.2 <b>Outlines</b> the theoretical physics concept(s) relevant to the research question	3.1.2 <b>Selects</b> materials and measuring equipment <b>appropriate</b> to the task	3.2.2 <b>Suggests</b> repeating the data collection	3.3.2 <b>Identifies</b> other significant variables that could influence the dependent variable	4.1.2 <b>Identifies</b> relevant safety issues	Students at this level can design an inquiry to explore the relationship between two variables.  Students consider factors that might affect the validity of the data collected. They show awareness of safety issues.
Remembering	1.1.1 <b>States</b> a research question or the title of the experiment	2.1.1 <b>States</b> a context for the inquiry	2.2.1 <b>States</b> a relevant area of physics	3.1.1 <b>Writes</b> instructions	3.2.1 <b>Identifies</b> the values for the independent variable	3.3.1 <b>States</b> one variable to control	4.1.1 <b>States</b> generic safety procedures	Students at this level can plan an inquiry to answer a physics-related question.
	1.1.0 Insufficient evidence	2.1.0 Insufficient evidence	2.2.0 Insufficient evidence	3.1.0 Insufficient evidence	3.2.0 Insufficient evidence	3.3.0 Insufficient evidence	4.1.0 Insufficient evidence	Insufficient evidence
Indicative behaviours	1.1 Formulates a research question	2.1 Justifies the inquiry	2.2 Establishes the scientific context	3.1 Designs a logical method	3.2 Plans to collect sufficient data	3.3 Plans to collect relevant data	4.1 Plans a safe method	LEVEL STATEMENTS
	1 Topic and research question	2 Background information		3 Methodology			4 Safety	

The following table is another example of a developmental rubric designed by the theory of knowledge (TOK) teachers at Międzynarodowe Szkoły Paderewski w Lublinie, in Poland. Adapted from the summative assessment criteria, it is used in peer- and self-assessments to support students to generate insights on their learning and provide specific and constructive feedback to one another.

Subject	Example	IB World School
<b>Theory of knowledge</b>	<i>Assessment rubric for formative purpose: Peer and self-assessment</i>	Międzynarodowe Szkoły Paderewski w Lublinie, Poland

## Introduction

The last section discussed how to engage learners and teachers in quality feedback and what makes feedback effective (“What makes feedback effective and how do we deliver meaningful feedback?”). However, the quality of feedback is only as good as the design of the assessment activity itself. Hence, this section focuses on how to design meaningful assessments that support quality inferences and rich feedback for both learners and teachers.

## Universally-designed assessments: Varied and multimodal

To engage learners with rich formative feedback, classroom assessment needs to be varied and multimodal, and it can then reflect the diversity of the knowledge, skills and abilities that DP learners develop throughout their learning experience. Therefore, DP teachers use a wide range of tasks and activities in IB classrooms to elicit evidence of student learning.

Effective formative assessment designs consider that every learner is unique and that each one has his or her own backgrounds, strengths, interests and areas to improve. Therefore, using a wide range of diverse assessments is a way to ensure that classroom assessments are fair and accessible and provide ample opportunities for students with different strengths to demonstrate their understanding and receive quality feedback.

These assessment tasks may be in written forms, such as an essay or a blog; in visual forms, for instance, a graph or a concept map; in verbal forms, such as an oral presentation or group discussion; or in numerical formats, such as a table or data analysis spreadsheet. Moreover, students are sometimes asked to perform what they have learned, in a play or practical activity for example. Classroom assessments could also be in mixed modalities, such as an exhibition or a video.

Figure 5 below gives an overview of the diversity of formative assessments. The list is not exhaustive, but it conveys the idea that formative assessments go beyond tests and written tasks and tap into students' understanding in different modalities.

Figure 5

The diversity of assessment types and modalities

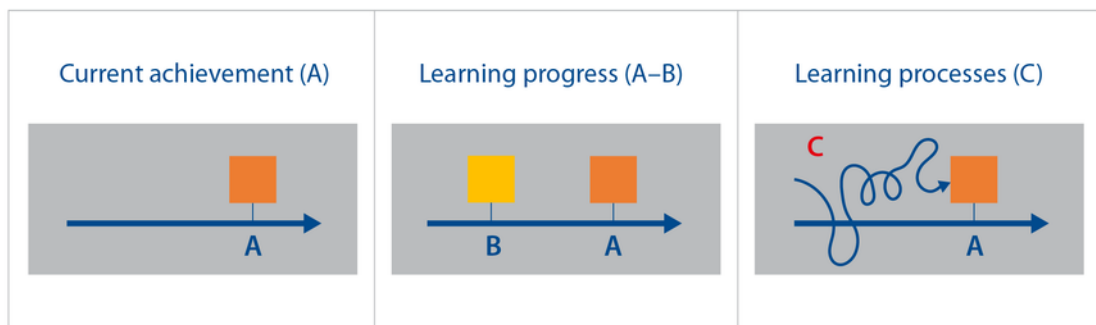
Assessment types	Modalities	Assessment types	
Results/solutions Data analysis	<b>Numerical</b>	Table Formula Problem-solving steps/process	
Thesis Report Dissertation Essay		<b>Written</b>	Hypothesis Annotated extracts Proposal Blog Journal Wiki
Diagram Graph	<b>Visual</b>		Web portfolio Note taking Illustration Highlighting text Concept or mind maps Sketch
← <b>Structured end-product</b>		→ <b>Fluid Processing</b>	
Presentation Podcast Reading scripts	<b>Verbal</b>	Viva Interview Conversation Debate Group discussion	
Performance Play Product portfolio Video Building models		<b>Performative</b>	Demonstration Rehearsals Process portfolio Questions Strategies
Lab experiment Theatre Song Exhibition Representations	<b>Hybrid</b>		Improvisation Practical activity Plan Steps Event Observations Explanations Activity

## Feedback that reflects current achievement, learning progress or process

Summative assessment is interested in finding out the level of achievement or competency of students; therefore, it focuses on what the students can do at the time of assessment. In formative assessment, the focus is on understanding what kinds of actions can be taken next to best help the students to learn. Teachers and learners are interested in formative feedback that not only provides information on current achievement but also reflects on the learning progress or processes of the students, such as the steps or strategies taken to approach a problem.

Figure 6

*What aspects of student learning do we want to provide feedback on?*

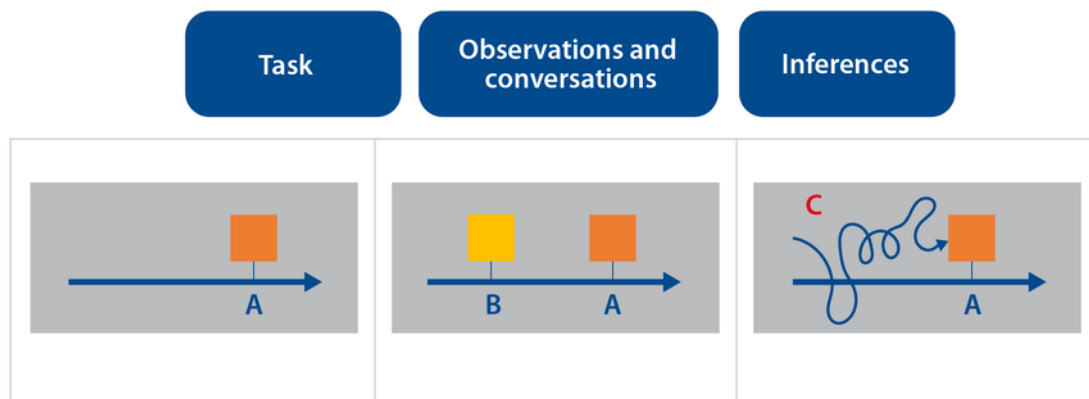


What is the best way to assess learning progress or learning process? Considering that effective formative assessments consist of three important parts: tasks, conversations and observations, and inferences, **good assessment designs first requires anticipation of the types of inferences made at the planning stage of assessment.** This means needing to anticipate the kind of feedback we want to share with learners and how to interpret the assessment outcomes before the actual assessment. This would then provide much clearer directions on how to design the task, what to observe during the learning process and what type of conversations to engage in with learners.

Let us look at an example of effective practices from DP mathematics. First, figure 7 is a reminder of the essence of formative assessment.

Figure 7

Three key elements of formative assessment



## Examples of effective practices: A task to examine both competency and learning processes

### Subject: DP mathematics, analysis and approaches

#### Assessment: What is the volume of this lemon?

In this formative assessment, the DP teacher designed a group experiment where students need to collaborate with their peers and find out as many ways as possible to measure accurately the volume of a lemon.

This was an open-ended task, where students had many different ways to approach an answer. In the planning stage of the assessment, the DP teacher anticipated the questions and inferences shown in the table below and incorporated them into the design of the assessment.

What do we want to assess? Key constructs/inferences	How do we assess this? Tasks, observations and conversations
Mathematical concepts	How well can students apply the mathematical concepts and content learned in class to solve a novel problem collaboratively?
Learning processes (Students' ability to collaborate in a problem-solving task)	Were students able to: <ul style="list-style-type: none"> <li>• begin work promptly</li> <li>• participate in group discussions</li> <li>• listen actively to group members</li> <li>• ask questions for clarification</li> <li>• build on what is being discussed and contribute</li> <li>• stay on task</li> <li>• approach problems creatively and be open to different approaches?</li> </ul>

In summative assessment, what is submitted to the teacher for marking or for written feedback is a structured end-product (the left-hand side of figure 5). Usually, it is not possible for the teacher to provide

feedback on the students' progress or the learning strategies or approaches taken just by looking at these learning products.

Nonetheless, in formative assessment, there is much more space to design a wide range of assessments, including those that tap into the adaptive learning processes of students (right-hand side of graph). In these activities, teachers not only focus on the answers produced by students but can use the opportunity to understand how students approach a problem, how they set learning goals or modify their learning strategies, how they respond to feedback and interactions, and how they reflect on feedback and improve their work. Hence, compared with mock exams and practice questions that are replicas of summative assessments, a wide range of assessment activities that include tasks, conversations and observations would provide much richer and informative feedback on students' learning progress and processes.

The table below provides some further examples of how assessment tasks may differ depending on whether the purpose of assessment is to measure current achievement or to elicit evidence of the learning process.

<b>Assessments that elicit evidence of current achievement or level of competency</b>	<b>Assessments that elicit evidence of learning processes</b>
Product portfolio	Process portfolio
Reading or listening comprehension questions	Observations of students' note-taking skills and their skills in highlighting key messages in texts
Thesis	Hypothesis
Exhibition	A plan or outline
Presentation	Conversations of students planning a group presentation



## Aspects of validity in formative assessment and feedback

### Do the same principles apply to formative and summative assessments?

The concept of validity has been referred to widely in the design and use of quality assessments. The IB generally focuses on five aspects of validity: construct relevance and authenticity, manageability, fairness and inclusivity, reliability, and comparability (figure 8), often referred to as the validity chain (Crooks et al 1996).

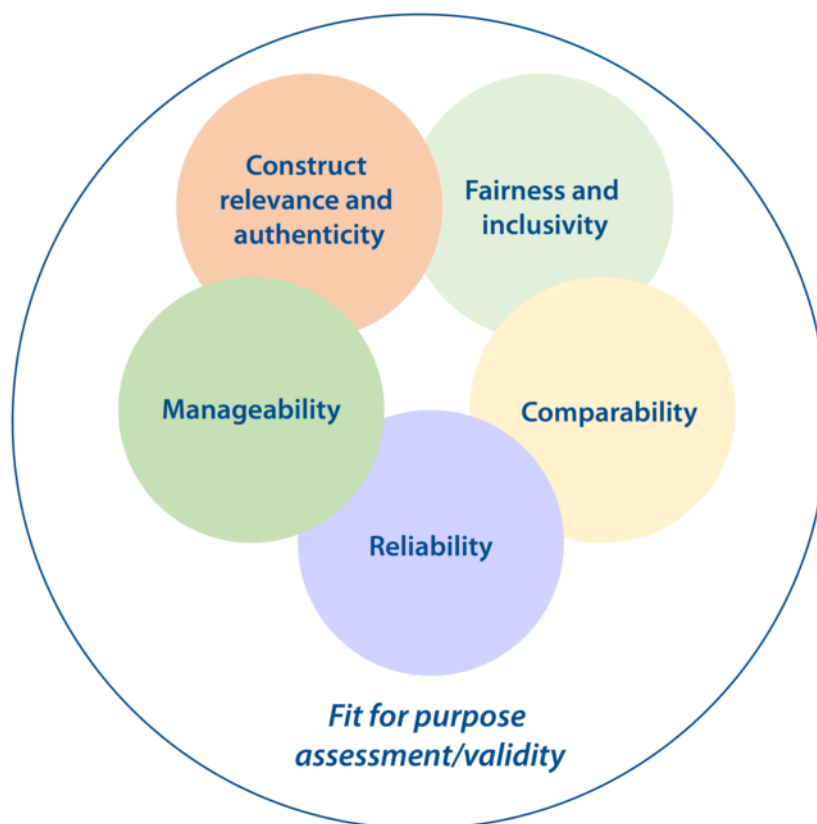
“These five elements are often in tension with each other. In considering whether an assessment is valid you need to consider the main purpose of the assessment to determine the relative importance of each aspect.”

“Each of the five elements in the chain are important in themselves but are not sufficient by themselves to make the assessment valid (that is, fit for purpose).”

—*Assessment principles and practices: Quality assessments in a digital age*

Figure 8

*Fit-for-purpose assessments: Aspects of validity in assessments*



From *Assessment principles and practices—Quality assessments in a digital age*, Figure 14 (2019)

Do the same principles apply to formative and summative assessments? How do teachers balance these aspects of validity to support meaningful feedback if we are not using marks or grades in formative assessments?

Although we are guided by the same principles when creating valid and meaningful formative and summative assessments, the different elements in the validity chain may hold different weights as we assess for different purposes. In formative assessment, we can focus more on designing construct relevant, fair and manageable assessments and removing assessment biases, while worrying a little less about reliability and comparability. This is because the quality of feedback is more important in formative assessments than how repeatable the grades are (if any grades were assigned). However, some level of comparability may still be helpful so that the conclusions drawn for student A and student B are meaningful.

## A supporting tool to balance aspects of validity and generate meaningful feedback

Figure 9 provides some examples of guiding questions to support DP teachers in making professional judgments about how to design meaningful assessments for formative feedback.

Figure 9

Some guidance on designing meaningful assessments

#### Authenticity and construct relevance

- **How well does the assessment reflect real-life situations?**
- **How do we design an assessment that accurately reflects the expected learning outcomes?**
- For example, if my intention is to assess “students’ ability to formulate a scientific hypothesis”:
  - what is the best assessment activity to provide feedback on this key construct?
  - how do I make sure that the activity is designed to assess only the primary skills (eg students’ ability to formulate a scientific hypothesis) and not other secondary skills (eg language abilities)?

#### Fairness and inclusivity

- **Do all learners have fair opportunities to participate in the assessment activity, to demonstrate their ability and to receive quality feedback?**
- **Does the assessment give an unintended advantage to a group of students without that being the point of the assessment?**
- For example, if my intention is to assess “students’ ability to conduct mathematical analysis”:
  - it is a fair assessment if it discriminates between student work that demonstrates better or poorer mathematical analyses
  - the assessment has an unintended bias if it discriminates between students with different digital skills or different levels of experience with spreadsheets.

#### Reliability

- **How confident are we that the student will get consistent feedback if the same type of assessment was repeated?**
- For example, would the student receive the same kind of feedback if the work was assessed by me, by another teacher or by a peer?
- Or, for instance, if two assessment tasks were designed to assess the same skills, would the student receive consistent feedback in both activities?

#### Manageability

- **How manageable is this assessment for the students and teachers in my school?**
- **When looking at the whole assessment experience across subjects and throughout the Diploma Programme, is the workload manageable and well distributed in time?**
- For example, the length of the assessments, the equipment or material required to complete the assessments, or the total number of assessments required
- **How can we make the assessments in our school more manageable for all learners?**

#### Comparability

- **How confident are we that the different assessment tasks are of comparable difficulty?**
- For example, if I have decided to give my students different options in an assessment task to make it more inclusive, allowing them to submit their work in the format of an essay, a presentation or a podcast.
  - Will the different options have similar levels of difficulty?
  - Will the different options allow me to provide equal quality feedback to all learners?

In some social or school contexts, DP teachers may be asked to assign a grade to students in formal formative assessments in addition to the qualitative feedback provided, or they may be required by state or district guidelines to use the outcomes in formative assessments to inform the predicted grades in their subjects. In these cases, extra attention may be paid to ensure that the grades used are reliable and

comparable. For further details, please refer to the IB publication *Assessment principles and practices—Quality assessments in a digital age* (2019).

How do DP teachers make use of these guiding questions to support their professional judgment around assessment designs and practices? Here are two examples, from DP physics and TOK.

## Example 1: DP physics

### Purpose of assessments

- To provide feedback on student ability to collaborate and conduct a simple group experiment on measuring half-life in beer foam
- For teachers and peers to provide feedback to one another on how to improve their ATL skills in sciences, especially their social, communication and research skills

### Possible assessment designs

The physics teacher considered two possible assessment designs for this task.

1. Ask students to collaborate in small groups and then submit individual lab reports for assessment. This option would have a high reliability and ensure that students are only assessed on their individual writing.
2. Ask students to submit a group lab report at the end of activity, and ask students to engage in peer-assessment and provide feedback to one another on their social, communication and research skills during the learning process.

In this scenario, the teacher chose the second task as she considered construct relevance more important than reliability in this formative task. The group report and peer-assessment format supported the purpose of making inferences and providing feedback to students on their ATL skills.

## Example 2: DP theory of knowledge

### Purpose of assessments

- To make inferences on how well students are able to explore the knowledge question, “Do moral truths exist?” by connecting this to the ways of knowing and the topics they have explored in class
- To understand how each individual learner is progressing in class and how comfortable he or she feels with the subject and the open-ended way of approaching knowledge questions
- To provide individual feedback to learners

### Possible assessment designs

The TOK teacher considered two possible assessment designs for this task.

1. To ask students to explore the knowledge question in a format of their own choice, for example, a five-minute podcast, an infographic, a short video or a mind map with drawings. As the learning product submitted by each student may look different, this task is of a low comparability. However, it allowed students to express their knowledge and understanding in a way that they feel most comfortable with, and this was thought to be a fairer and more inclusive assessment design to support the diverse learners in class.
2. To ask students to submit a short essay exploring the knowledge question, including a personal reflection on the ethical dilemma activity they engaged with in class. The teacher considered this task to have high comparability. She also thought that she could provide more reliable feedback on this task by using the summative assessment criteria.

On this occasion, the teacher chose the first option; she weighted fairness and inclusivity as more important than comparability and reliability in this task. The multimodality and student choices built in to the assessment design allowed the teacher to understand how well the students are progressing in terms of exploring the key concepts and open-ended way of approaching questions in TOK. Since the teacher's students are only in their first year of the DP, and most of the students are responding in their second

language, the teacher also wanted to minimize assessment biases by not focusing on students' essay writing skills.

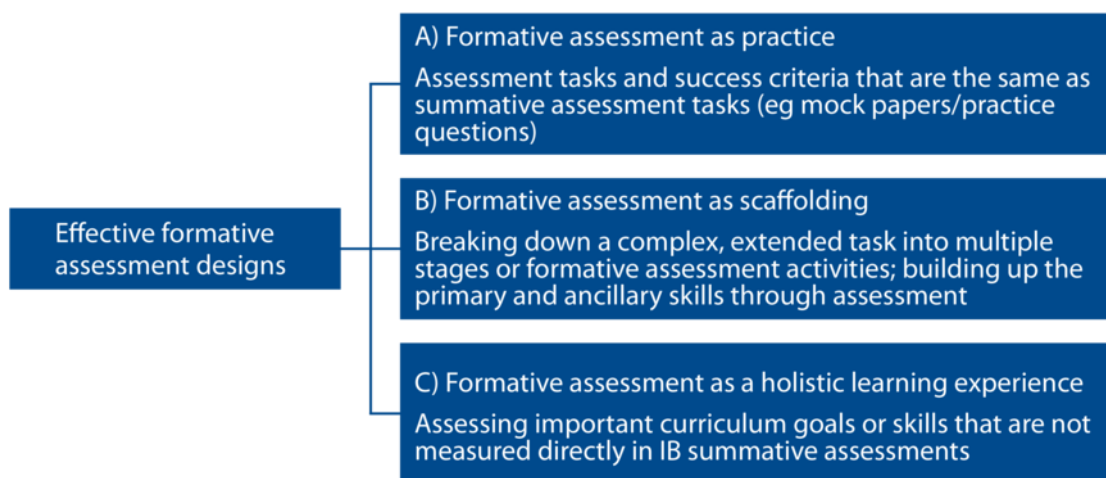
These examples show that assessment principles such as aspects of validity are not just abstract concepts that are discussed in theory. Many DP teachers who are fully eloquent in these skills often weight these principles implicitly when they make decisions about quality assessment and feedback in the classrooms. Having discussed the general principles that guide teachers to make decisions about assessment, the next sections go into more depth regarding the *what* and *how* to assess.

## Three main types of effective designs in DP classrooms

Formative assessment has different task design and success criteria depending on what teachers intend to measure (the test constructs). In conversations with IB educators and visits to DP classrooms, a wide variety of effective assessment designs have been identified in DP classrooms. These formative assessments can be broadly categorized into three groups (figure 10).

Figure 10

Three approaches to effective formative assessment design



### Formative assessment as practice

- Formative assessments as practice are a common type of formative assessments in DP classrooms. They measure the same test constructs and have the same task designs and success criteria as the IB summative assessments. These often occur in the format of practice questions, practice essays and mock papers, and are replicas of assessment components or questions in the high-stakes summative assessments.
- Formative assessments as practice are helpful in supporting students to know what is expected from them in the assessments, also known as “test-wiseness” (Baird et al 2016). The coaching of students to be test-wise was reported to increase assessment validity as it reduces construct irrelevant responses from students.
- The feedback given based on the success criteria of IB summative assessments could help students gauge “how far they are” from their individual learning objectives. However, it may not be the most effective in supporting students to know “how to get there” and to decide on what actions to take to improve their learning.
- Practice tasks and questions have their limitations as a formative assessment tool and should not be the only type of formative assessments occurring in DP classrooms.

For example, scaffolded tasks that break down a complex task into smaller steps by focusing on specific or multiple skills at a time are often better than practice tasks at providing fine-tuned and concrete feedback to learners.

Learners need experience with a wide range of assessment activities beyond practice questions to be able to transfer their understanding to new or unfamiliar contexts.

## Formative assessment as scaffolding

- Formative assessment as scaffolding is another type of formative assessment that breaks down a complex extended task into multiple stages or various formative assessment activities. These may best address the needs in a skills-based curriculum (see the section, [What do classroom assessments look like in a skills-based curriculum?](#)).
- The aim of formative assessments as scaffolding is to support teachers and learners to identify and assess the skills learners need to build up to reach their learning objectives in the subject and beyond, and to perform well in the summative assessments.
- Learners need to identify the ancillary skills that the tasks require but are not explicit in the learning or assessment objectives. For example, here are some ancillary skills identified by IB educators.
  - Knowing when to use a tool or technology
  - Organizing and presenting ideas and thoughts in writing
  - Understanding key words or command terms in questions
  - Using visuals to represent knowledge
- Teachers that design and use formative assessments as scaffolding often have a deep understanding of the learning objectives of the course of study and the strengths, learning progress and needs of their students so that they can scaffold the assessment activities and feedback responsively. These assessments cannot be simple replicas of summative assessment questions and tasks.
- Formative assessments as scaffolding are also the type of assessments that novice teachers may find more challenging and would like more support in assessment design and delivery. These may be more challenging in subjects where the high-stakes summative tasks are more open-ended and tap into complex skills that cannot be readily divided by content or broken down into smaller, problem-solving steps, for example, in some individuals and societies subjects or in TOK.
- In interviews with IB educators (in the pilot study), formative assessment tasks as scaffolding were reported to be slightly more straightforward in subjects such as mathematics, sciences and language acquisition.
- How do teachers identify the primary and ancillary skills in a curriculum, and how can teachers use formative assessments to scaffold these skills and track progress?

An example of formative assessment as scaffolding may look like the table below.

DP subject	Geography higher level
<b>Extended task</b>	Evaluate the costs and benefits of hosting an international event (eg the Olympics)
<b>Examples of formative assessment questions or tasks</b>  <b>(Note: Teachers and learners can approach these step by step, but also focus on isolated skills to elicit evidence of learning and provide feedback)</b>	<ul style="list-style-type: none"> <li>Evaluate different types of sources and information.</li> <li>Argue for choosing one source over the others or explain the reliability of your sources.</li> <li>List all the costs or list all the benefits.</li> <li>Compare and contrast the costs and benefits according to different aspects or factors (eg social, economic, environmental).</li> <li>Present your views in writing, presentation or video format.</li> <li>Consider ways to revise a piece of writing for a different audience.</li> </ul>

DP subject	Geography higher level
	<ul style="list-style-type: none"> <li>Consider ways to improve the introduction of a sample essay.</li> <li>Explain the strong and weak points of an argument.</li> </ul>
<b>Possible activities or formats of assessment</b>	Self-assessment, peer-review, group discussion, essay, concept or mind maps, debate, worksheet, interview, plan, learning journal, highlighting text, presentation, etc
<b>Student agency</b>	Providing choices: students deciding which task or skills to focus on, peer-assessment and self-assessment, co-constructing success criteria

## Formative assessment as a holistic learning experience

- Formative assessment as a holistic learning experience addresses important curriculum goals or skills beyond the summative assessments. These may be goals and skills that are hard to assess in current high-stakes assessments. Some DP educators would identify these as collaboration, creativity, attributes such as the learner profile, and other qualities that are important for students to develop.
- These elements that go beyond what is assessed directly in high-stakes summative assessments may sometimes be taken away from formative assessments, especially in a curriculum that is content-heavy and when learners and teachers are under time constraints to prepare for summative assessments.
- Attributes such as collaboration, creativity, or the learner profile attributes, are not addressed directly in the current high-stakes assessments. So the challenge is for IB teachers to design meaningful assessments that could address these qualities. For example, how do we assess collaboration, creative writing, prediction and decision-making skills, or students' abilities to interact in dialogues and solve a real-life problem?

The table below lists some of the skills reported by IB educators as essential for the learning objectives in DP subjects, but that are also sometimes hard to assess in high-stakes summative assessments or in school examinations.

Areas of assessment	Essential skills reported by IB educators that could be challenging to address in current summative assessments
<b>Problem-solving (novel problems) or inquiry</b>	<ul style="list-style-type: none"> <li>Intuitive problem-solving</li> <li>Ability to work with the unknown</li> <li>Project-based activity</li> <li>Open-ended scientific tasks</li> <li>Design thinking</li> <li>Mathematical investigation and inquiry</li> <li>Sustained reasoning, sustained problem-solving</li> </ul>
<b>Collaboration</b>	<ul style="list-style-type: none"> <li>Relating to peers and supporting each other in a learning task</li> <li>Group work</li> <li>Collaborative problem-solving</li> <li>Communication skills</li> <li>Peer-assessment or peer review</li> </ul>
<b>Other higher-order cognitive skills</b>	<ul style="list-style-type: none"> <li>Higher-order critical-thinking skills</li> <li>Decision-making</li> </ul>



Areas of assessment	Essential skills reported by IB educators that could be challenging to address in current summative assessments
	<ul style="list-style-type: none"> <li>• Forecasting and projecting trends</li> <li>• Linkage to TOK (to reflect on what students know and how they had learned)</li> <li>• Ability to synthesize and deconstruct a complex task</li> <li>• Analysing and production of non-written or multimedia tasks</li> <li>• Visual-spatial skills (the ability to comprehend and express ideas through visual images and visual-spatial experiences)</li> <li>• Asking good questions</li> <li>• Systems thinking</li> </ul>
<b>Learner attributes</b>	<ul style="list-style-type: none"> <li>• IB learner profile, such as               <ul style="list-style-type: none"> <li>Reflective thinking</li> <li>Risk-taking</li> <li>Principled</li> <li>Open-mindedness</li> </ul> </li> <li>• Resilience/learning from failures</li> <li>• Student agency</li> </ul>
<b>Creativity</b>	<ul style="list-style-type: none"> <li>• Originality</li> <li>• Curiosity/inquisitiveness</li> <li>• Creative writing</li> <li>• Adaptability</li> </ul>
<b>Learning process and progress</b>	<ul style="list-style-type: none"> <li>• Interactive dialogue/dynamic debate</li> <li>• Ability to interact in learning activity and respond to feedback from teachers and peers</li> <li>• Thinking/learning process</li> <li>• Progression of thinking</li> <li>• Individual progress</li> <li>• Skills that are transferrable across content and topics</li> <li>• Approaches to learning (ATL)</li> </ul>
<b>Digital literacy/tools and technology</b>	<ul style="list-style-type: none"> <li>• Understanding when and when not to use technology</li> <li>• Designing tools and equipment</li> <li>• Digital literacy</li> <li>• Searching for relevant information on the internet; differentiating between what is important or useful</li> </ul>

## Is “teaching to the test” quality formative assessment?

### Teaching to the test

There is a fourth type of assessment that is often confused with the tasks in part A of figure 10 (formative assessment as practice) and that should not be considered good formative assessment. It relates to teaching to the test whereby students and teachers resort to using templates, rote memorization, question anticipation and pre-prepared or pre-rehearsed responses to maximize grades in high-stakes assessments, instead of training the higher-order cognitive skills required to perform in those novel tasks or problems.

This type of drilling on test content is reported to undermine test validity. One of the reasons for this is that teachers and learners may “narrow the taught curriculum” as they use question anticipation and leave out certain topics that they believe will not be tested in the external examinations (Baird et al 2016; Elwood et al 2017). Therefore, students risk resorting to more surface learning techniques that may take them away from the full learning experience and the personal fulfilment derived from deep engagement with learning materials.

This type of rote memorization goes against the IB’s pedagogical philosophy. Moreover, the students may, ironically, perform less well in final summative assessments as a result of such an approach because they fail to develop the higher-order thinking skills that are needed to meet the full requirements of the criteria. For example, research has shown that predictive “question spotting” used by teachers or learners to narrow the taught curriculum was related to lower relative progress in student performance (Greatorex, Malacova 2006).

## Inquiry-based and concept-based curriculum

One of the approaches to teaching in the IB is **teaching focused on conceptual understanding**. What does formative assessment look like in a classroom where learning and teaching is conceptually driven? How do teachers assess the development of conceptual understanding in students?

As learning and teaching in DP classrooms is increasingly driven by big questions and organized around interconnected concepts, both subject-specific and transdisciplinary ones, this shapes the way teachers design and approach formative assessments. Instead of designing content-based assessments or end-of-unit tests to assess formatively facts and knowledge, IB teachers often need to design assessments that focus on students' increasing complexity and depth in conceptual understanding. This also means assessing students' abilities to organize and connect knowledge and concepts and apply them to new or unfamiliar contexts.

# What are concepts and how do we assess conceptual understanding?

## What are concepts?

There are slightly different ways of understanding concepts and conceptual understanding among educators. For discussion purposes, in this publication concepts are broadly defined as abstractions or the “big ideas”. Concepts are mental representations of categories of objects, events or other entities (Jonassen 2006).

## How do we assess conceptual understanding?

The IB has published a research report on concept-based teaching and learning (Medwell et al 2019). Some of the main ideas in the report and the wider literature on assessing conceptual understanding in students are as follows.

- **Authenticity**

To develop students’ conceptual understanding, the IB uses “authentic assessment” as much as possible (Wiggins 1993). This is done by engaging students in worthy questions and real-life problems so that they can apply their knowledge and concepts in a creative and effective way.

- **Application**

The IB does not assess conceptual understanding simply by asking for the definition of concepts. It is more important for students to apply the concepts in different contexts.

- **Cumulative learning**

The development of conceptual understanding is like an upward spiral rather than a linear stage progression. As learners revisit concepts in various contexts, they demonstrate increased breadth, depth and complexity in their understanding (Chadwick 2009).

The goal in assessing conceptual understanding for formative insights is to evidence this increasing breadth, depth and complexity in understanding and to provide effective feedback to move learning forward.

- **Students can be assessed on how well they have understood a concept by being asked to:**

identify examples of a concept

organize and reflect on this and provide counter-examples

develop generalizations

apply developed generalizations to past and future knowledge and make accommodations when needed (VanTassel-Baska, Wood 2010).

- **Applying understanding in new or unfamiliar contexts (sometimes known as transfer skills)**

Using case studies or content that is slightly different from the examples taught in class but that the students can relate to.

This could help to assess students’ ability to apply the concepts in new or unfamiliar contexts, which is the primary aim of developing their conceptual understanding.

Understanding is demonstrated when students can transfer their learning autonomously.

- **Social construction of meaning**

Concept-based models are more powerful when students are engaged in collaborative group work and conversations about the application of concepts.

Design group or collaborative assessment activities so that students can work together to generate new ideas and reframe their understanding about concepts.

## Practical examples from DP classrooms

Some additional tools shown to be helpful (Medwell et al 2019) for assessing conceptual understanding include the following.

- Semi-structured interviews
- Guiding questions
- Open-ended essays
- Think-aloud problem-solving
- Graphic organizers
- Concept maps
- Know-wonder-learned charts, commonly called KWL charts, guide learners to express what they already know on a topic, to list questions they want to know, and demonstrate what they have learned
- Word clouds

Using graphic organizers and concept maps can be helpful for the following reasons.

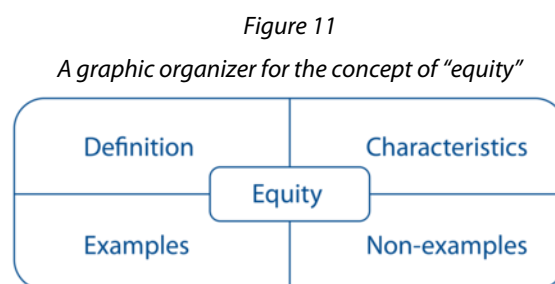
- Asking students to express their ideas graphically can provide a powerful window into their thought processes and understandings, because thinking is not linear, it is more like a map.
- Inviting students to express their ideas in the form of graphic organizers or concept maps aligns well with findings in cognitive science that our minds organize ideas through **schemas** (Piaget, Cook 1952).
- Assessing visually could sometimes be more powerful than assessing through texts as teachers can see how well students are able to connect various concepts and knowledge, and display increased complexity in their conceptual understanding.

### Effective formative assessment practices: Example 1

#### Using a Frayer model graphic organizer to assess conceptual understanding

**Subject:** DP economics (see the *Economics teacher support material*)

**Concept:** Equity



For formative assessment, students are introduced to the economic concept of “equity” in this class. The DP teacher uses the Frayer model graphic organizer (figure 11) to assess students’ conceptual understanding as they conduct inquiry and explore various case studies. The graphic organizer is an effective way of helping students make their thinking processes explicit and present them visually. Besides asking students to define the concept “equity” and describing its characteristics, which requires lower-level cognitive demands, students are also asked to explore examples and non-examples of “equity” as they encounter various case studies and scenarios inside and beyond the classroom.

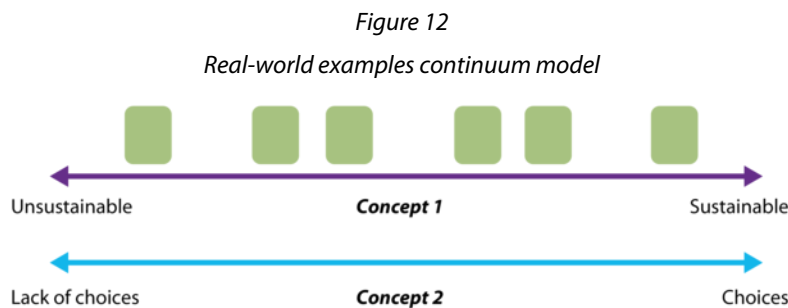
Students are encouraged to work in small groups in this activity. This is also a great exercise to explore possible misconceptions that the students may be holding, especially through the examples and counter-examples that they raise.

## Effective formative assessment practices: Example 2

### Using a card sorting activity and group discussions to assess conceptual understanding

**Subject:** DP economics (see the *Economics teacher support material*)

**Concepts:** Sustainability, choices



This is a card sorting activity to see how well students can apply a concept to understand and compare real-world examples. The idea is to place a number of cards that record different real-world examples along a continuum of a key concept—for example, cards for national or local economies, different types of firms, economic policies or specific case studies relating to the key concept (figure 12). In this example, students sort the cards according to the degree of “sustainability”—from the most sustainable practices to the least. Students discuss with each other and assess the extent to which each concept applies to the examples. Students’ conceptual understanding can be assessed as follows.

#### Observations

- Which are the concepts or examples that students demonstrate deep understanding of? Do students struggle with particular concepts or examples? What can this tell us about their conceptual understanding? Are there significant misconceptions or gaps in understanding?

#### Conversations

- The order of the cards is not the most important element in this open-ended task. There are different possible ways to arrange the cards along the continuum.
  - Are students able to understand and express the nuances in this activity?
  - Are students able to compare the results and discuss when their ideas differ?
  - Could students support their arguments with rationales?
  - Are students able to identify what is unexpected and find reasons for these?
- Once students have sorted the cards with multiple concepts, such as “sustainability”, “intervention” and “choice”, the following can be explored:
  - Can students draw connections between the concepts?
  - Are students able to draw conclusions and generalizations about the concepts and their application in real-life examples?
  - Can students identify new questions that emerged as they explored the connections between concepts?

## Effective formative assessment practices: Example 3

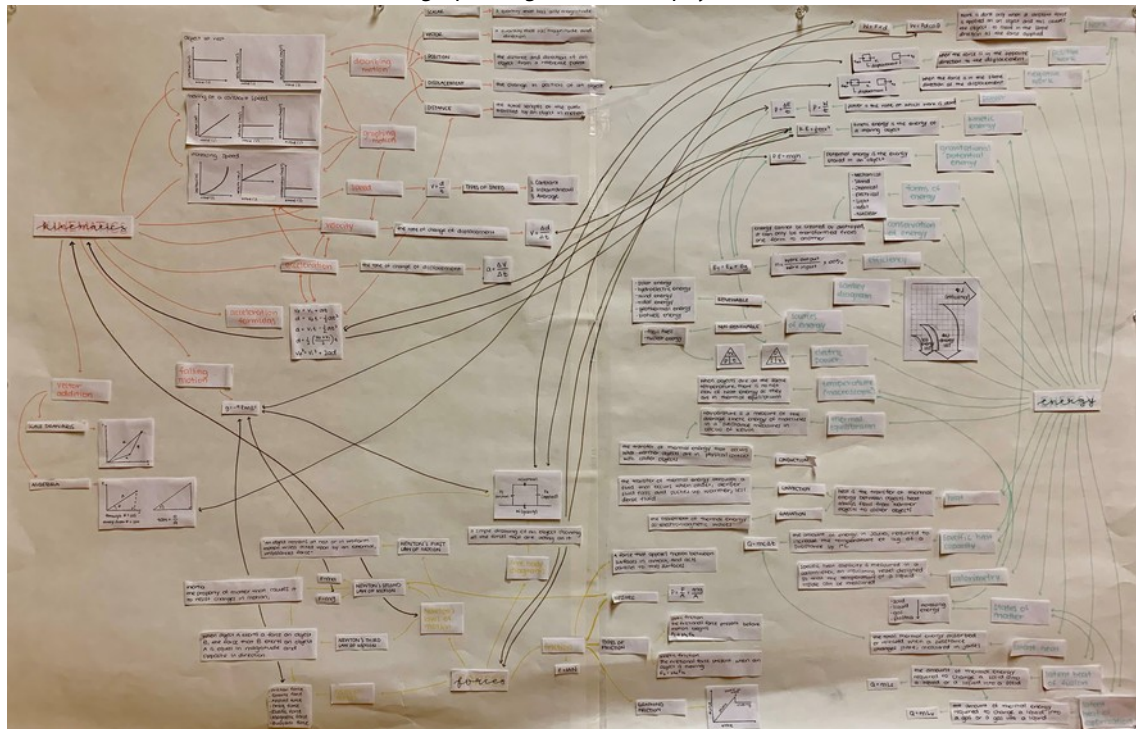
### Using graphic organizers to assess conceptual understanding

Subject: DP physics

Concept: High-level concepts and micro-topics

Figure 13

A graphic organizer for DP physics



Example shared by Dr Jennifer Elliott, IB Diploma Coordinator/Physics Teacher, Toronto French School—Canada's International School

In this formative assessment activity, students draw and build this concept map collaboratively as they learn new concepts and content throughout the DP physics course. As illustrated from the map (figure 13), students not only tried to connect the various over-arching and high-level concepts in the subject, such as “energy” and “kinetics”, but also explored the connections between the various micro-topics.

This is an effective example to assess students’ development of conceptual understanding in a subject. As learning progresses, students can see that they demonstrate increasing knowledge of the concepts, content and topics, and most importantly that they can draw more and more connections between the various concepts, as illustrated by the complex and dynamic arrows and lines connecting them. Sometimes, students also encounter important learning moments when they realize they need to redraw or reframe their maps. This is similar to the way our mental maps or schematic networks work in our minds (Ültanır 2012).

As teachers and students explore the concept map together, they can identify whether there are significant gaps or misconceptions in learning and provide helpful feedback to one another to inform the next steps in learning and teaching.

Example of effective practice	Targeted skills
<ul style="list-style-type: none"> <li>• <i>DP physics: Assessment rubric and activity—Converting sketch notes and learning conversation</i></li> </ul>	<ul style="list-style-type: none"> <li>• Conceptual understanding in physics</li> <li>• Ability to draw connections between multiple key concepts</li> <li>• Application and analysis of key concepts</li> <li>• Communication skills</li> </ul>

## Further examples

### Visible thinking routines

A research-based conceptual framework has been developed by Harvard Graduate School of Education (<https://pz.harvard.edu/projects/visible-thinking>).

- The purpose of the framework is to make thinking processes visible.
- The framework cultivates in learners the habit of mind to become critical, creative and deep thinkers.
- The three key elements of the framework—thinking routines, documentation of student thinking, and reflective professional practice—align well with effective formative assessment practices.
- The framework encourages collaborative thinking.

Some examples that are helpful for assessing conceptual understanding are shown in the table below.

<b>Connect, extend, challenge</b> ( <a href="https://pz.harvard.edu/sites/default/files/Connect%20Extend%20Challenge_0.pdf">https://pz.harvard.edu/sites/default/files/Connect%20Extend%20Challenge_0.pdf</a> )	A routine for drawing connections between new ideas and prior knowledge
<b>3-2-1 bridge</b> ( <a href="https://pz.harvard.edu/sites/default/files/3-2-1%20Bridge_0.pdf">https://pz.harvard.edu/sites/default/files/3-2-1%20Bridge_0.pdf</a> )	A routine for organizing understanding of a topic through concept mapping
<b>Generate, sort, connect, elaborate</b> ( <a href="https://pz.harvard.edu/sites/default/files/Generate-Sort-Connect-Elaborate_0.pdf">https://pz.harvard.edu/sites/default/files/Generate-Sort-Connect-Elaborate_0.pdf</a> )	A routine for organizing understanding of a topic through concept mapping

More thinking routines can be found on *Project Zero's Thinking Routine Toolbox* (<https://pz.harvard.edu/projects/visible-thinking>).

## Taxonomies to assess conceptual understanding

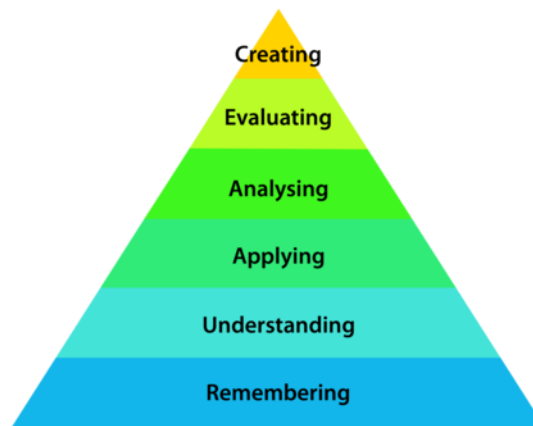
Taxonomies are another powerful method to assess learners' conceptual understanding.

### Revised Bloom's taxonomy

- Many teachers are familiar with the revised Bloom's taxonomy (figure 14; Anderson, Bloom 2001), which is used to assess increasing demands of cognitive skills.
- The command terms used in DP high-stakes assessments are also connected to the different levels of Bloom's taxonomy.



Figure 14  
Revised Bloom's taxonomy on Cognitive Processes



Adapted from Anderson and Bloom (2001: 268)

## The Structure of the Observed Learning Outcome taxonomy

- The Structure of the Observed Learning Outcome (SOLO) taxonomy is a powerful tool to support the assessment of conceptual understanding in students (figure 15; Biggs, Collis 2014).
- The aim of SOLO is to provide a systematic way to describe how a learner's performance grows in cognitive complexity as he or she gains proficiency in academic tasks.
- The SOLO taxonomy is helpful in assessing the increasing quantity and quality of thoughts that students demonstrate (Hattie, Brown 2004).

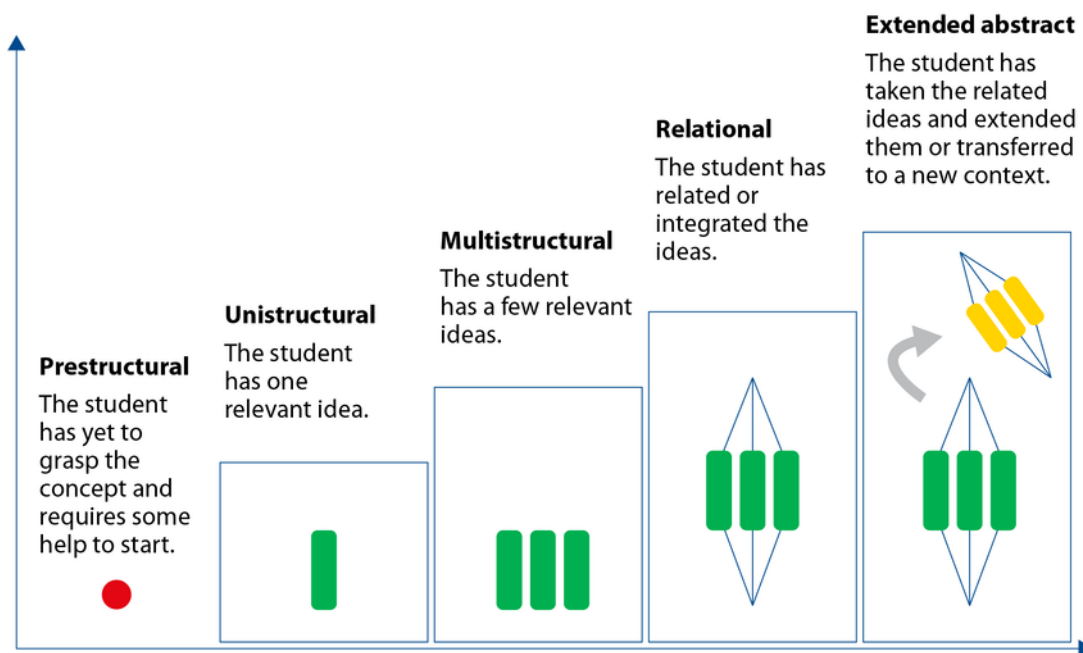
SOLO supports the assessment of surface and deep cognitive processes, which are both needed to perform in learning tasks.

SOLO shows how students may be demonstrating different levels of understanding and cognitive complexity on the same concept or topic, and therefore supports teachers and students in giving feedback to one another.

The same student may demonstrate distinct levels of understanding on different concepts, sub-concepts and applications. For example, a DP physics student may demonstrate relational understanding in "conservation of momentum", while he or she is still at unistructural understanding in "conservation of energy".

- While the use of an analytic markscheme can help to assess how much content and knowledge students can analyse, and how well they can define and describe concepts, the SOLO taxonomy may be more helpful in assessing how well students are able to make connections between ideas and transfer the concepts to new contexts.
- Taxonomies such as the SOLO taxonomy can be used to develop assessment activities that measure students' underlying understanding in areas of knowledge.

Figure 15  
SOLO taxonomy



Adapted from Belshaw (2017), Biggs and Collis (2014: 24–25)

## Assessing through extended conversations

- One of the key elements in formative assessments is using “conversations and observations” because we are interested in not only students’ answers to a question but also their thought processes and how they derived those answers.
- Using extended conversations is a good way to assess conceptual understanding and to support students to integrate knowledge.
- Through the use of extended conversations, questioning and feedback among teachers and peers, students can:
  - extend and apply their knowledge of concepts
  - make connections between concepts and between what they have learned and prior knowledge
  - reframe existing conceptual understanding
  - sometimes transform understanding, especially when they learn a “threshold concept” that completely shifts their understanding. According to Meyer and Land (2003: 1), threshold concepts open up “a new and previously inaccessible way of thinking about something ... without which the learner cannot progress”
  - engage in peer-assessment and feedback to expand and reconstruct their conceptual understanding through collaborative thinking and learning.

## Using multimodal assessments

- Another way to support the development of conceptual understanding is to encourage students to represent their knowledge in multiple ways (Bennett, Gitomer 2009). The IB therefore encourages the use of multimodal assessments to assess conceptual understanding in formative assessments.
- IB teachers design a wide variety of assessment tasks or activities that ask students to demonstrate their understanding in written, visual, verbal, numerical or performative action forms. This ensures that students with different strengths and learning needs can have fair opportunities to demonstrate their conceptual understanding.

### Effective formative assessment practices: Example 4

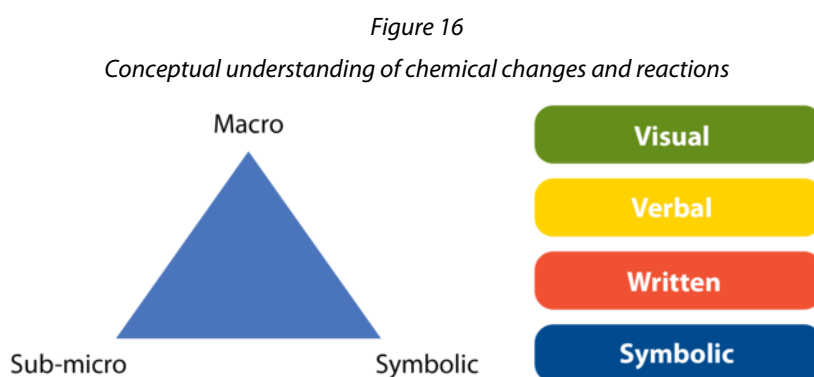
#### Using multimodal assessments to assess conceptual understanding

**Subject:** DP chemistry

**Concept:** Chemical changes and reactions at macro, sub-micro and symbolic levels

A DP chemistry teacher has shared her experience in using the Johnstone's triangle (figure 16; Taber 2013) and multimodal assessments to support students to develop and demonstrate their conceptual understanding of chemical changes and reactions.

In the formative assessment exercises, the teacher invites students to engage in peer-assessment and to demonstrate their understanding in visual, written, verbal and symbolic ways. She explains that when a student is able to explain verbally or in writing a chemical reaction at the macro level (for example changes in substances), to draw the sub-micro (molecular) changes visually, and then to represent their understanding in symbolic and numerical ways (chemistry symbols and formula), she is confident that the student has fully grasped the concepts and is ready to move on in his or her learning.



## Assessing possible misconceptions

- The distinction between formative and summative assessments is in the purposes of assessment and types of inferences drawn from assessment outcomes (Black, William 2018).
- In summative assessments, the focus is on the competency or level of achievement of a student, therefore “what the student can do after a course or unit of study”.
- In formative assessments, besides “what the student can do at the time of assessment”, we are also interested in “why a student did something and how to help to support them to move learning forward”, therefore trying to understand the thinking processes that led to the responses.

- One effective way to identify gaps in understanding is to analyse patterns of responses in students, for example, supporting learners to understand whether an unexpected response was due to an error, a careless slip, misconceptions about a key concept, or a lack of understanding (Bennett 2011).
- Analysing patterns of responses by students could support teachers to identify individual needs and adjust instruction accordingly, such as knowing:
  - which students could act as learning resources for their peers
  - which students may need minimal feedback or prompts
  - when it is appropriate to reteach a certain procedural or conceptual knowledge
  - when there may be a significant misconception and that students require more support to engineer a deeper cognitive shift (Bennett 2011).
- The use of short-response questions as effective formative assessment tools has been well documented in literature (Christodoulou 2017; Wiliam 2011). These may be more common in subjects such as mathematics and sciences. The use of multiple-choice questions or “true or false” questions is suggested by many teachers as being useful for formative assessment, especially when the distractors are responses that students may choose if they have a fundamental misconception about the topic.
- While these selected response questions were thought to be quick assessments that were easy to administer, effective questioning requires prior thinking about where students might feel challenged or confused about a topic.
- Often in IB classrooms, teachers use a combination of short-response questions, conversations and observations to assess student conceptual understanding, as illustrated in example 5. While the design and use of online quizzes or multiple-choice questions is a quick way to assess students’ concepts and misconceptions on a topic, the follow-up questions that students and teachers engage in is a remarkable way to evaluate students’ reasoning behind unexpected responses. Many teachers ask students to engage in discussions when a group of students demonstrates very different responses or approaches to a question. This is also a good way to identify students whose response did not necessarily match the markscheme but who demonstrated exceptional understanding of the concepts and are able to “think outside the box”.
- While there are many examples in literature on how to assess possible misconceptions using short-response questions, it is also possible to identify misconceptions through open-ended problem-solving or extended response tasks. A promising way to do this is through extended conversations and observations, as illustrated in the example 6.

## Effective formative assessment practices: Example 5

### Identifying possible misconceptions in learning

**Subject:** DP chemistry

**Activity:** Draw an example of a covalent bond and highlight where the bonding is

In the pilot study, a DP chemistry teacher shared that students who demonstrated deep understanding of the concept would draw a covalent bond that “has a shared pair of electrons under the influence of two nuclear isotopes that are positive, and that’s the bond”. However, there are students that only demonstrated partial understanding by drawing a shared pair of electrons. “Even as simple as that, there’s a misconception there. The actual bond is the positive nucleus holding your negative electrons and positive and negative attract.” Identifying students’ possible misconceptions at this stage of learning is helpful because “if they believe it’s only a shared pair of electrons, this may bring a lot of trouble later on as they learn further concepts”.

## Effective formative assessment practices: Example 6

### Identifying possible misconceptions in an open-ended task

**Subject:** DP history

**Group discussion:** *Who was to blame for the cold war?*

A history teacher in the pilot study shared that the common misconceptions that some students have in the subject is believing that “History is a sequence of facts and events in the past” and that “There is only one history. There is a right or wrong way to approach a question”. The teacher noticed this understanding was more common among students who came from a different curriculum and were used to the subject being taught in a very different way.

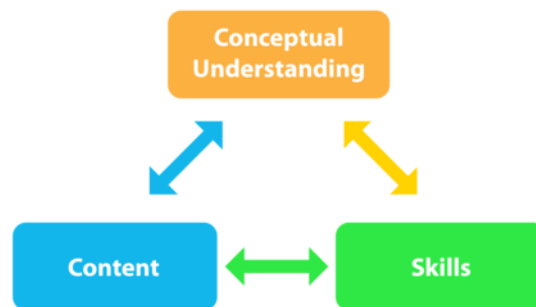
In this class, the teacher asked the students to engage in a group discussion on “Who was to blame for the cold war?”. Students were assessed on their ability to argue and interpret based on multiple perspectives, to analyse critically multimedia texts, and to adjust their ideas as they were presented with new facts and arguments. Through the conversations and observations, the teacher was able to identify where students might be stuck or have misconceptions about the subject. This is a simple yet excellent activity to demonstrate how it is possible to assess for possible misconceptions in an open-ended complex task and not only in closed-ended short-response questions.

## The development of skills within a concept-based curriculum

There is an increased recognition in education that learning is not just about the acquisition of facts and knowledge but also about practising and applying skills. In educational psychology, these two concepts are sometimes referred to as **declarative knowledge**, the “knowing-that”, and **procedural knowledge**, “the knowing-how”, respectively. The purpose of assessment is to find out if students “know or understand” something, but also whether they can apply the high-order cognitive skills in real-life or authentic contexts. The relevance of a skills-based curriculum has also gained importance as educators talk about how to integrate 21st-century skills, such as critical-thinking skills, creativity or collaboration skills into learning, teaching and assessments. These are usually skills that are considered essential to the future success and personal development of students (Dede 2010). In recent DP curricular reviews, many DP subjects have shifted away from content-rich curriculum to one that emphasizes the concurrent development of skills, conceptual understanding and content (figure 17).

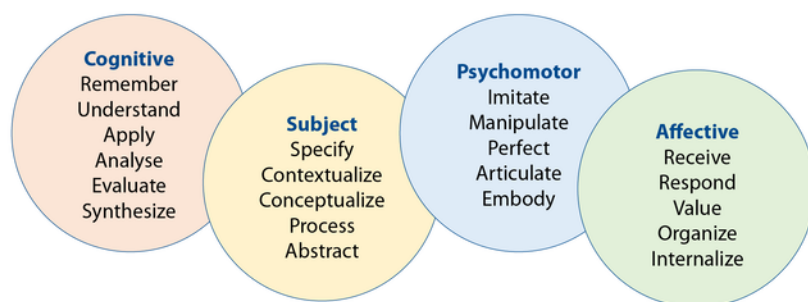
Figure 17

Three important and interacting elements in the DP curriculum



The suggestion to differentiate between concepts, content and skills is consistent with research in cognitive science and educational psychology. Often, a distinction is made between discipline knowledge—for example, the ability to contextualize and conceptualize knowledge—and cognitive knowledge, which encompasses the more transdisciplinary higher-order thinking skills, such as the ability to analyse, evaluate and synthesize (figure 18; Atkinson 2013). The challenge in formative assessment is then to identify what these skills are and how evidence of their development can be elicited along the learning journey in DP classrooms.

Figure 18  
Intended learning outcomes by Atkinson



Adapted from Atkinson (2013: 6)

Here are some key questions to explore when thinking about assessments in a skills-based curriculum.

- What do formative assessments look like in a curriculum that emphasizes the concurrent development of concepts, content and skills?
- What are the core skills that IB educators considered important, to be developed in the DP subjects and at a transdisciplinary level?
- What are the learning progressions of the core skills that would support multiple entry points of learning and teaching in the course of study as well as the formative assessments of individual learning needs and progress?
- How could members of the IB community support each other to reimagine and design these types of responsive formative assessments centred around the development of skills and conceptual understanding?

There are various possibilities in response to the questions above. In interviews with IB members of staff, many pointed out that research skills, critical reading and critical writing skills, and the skills to engage with helpful technology and tools, are examples of the core skills that students are expected to develop and have in their DP learning journey. These skills are mentioned for their importance in supporting students to engage with new or unfamiliar content, which is crucial for student agency in a concept-driven course of study. The table below shows some examples of the core skills.

Examples of skills to be developed	Possible formative assessments: Sub-skills or steps to build up core skills
<b>Skills in conducting and presenting research</b>	<ul style="list-style-type: none"> <li>• Asking good research questions</li> <li>• Developing and evaluating ways to collect evidence</li> <li>• Research methodology</li> <li>• Analysing data or examining the evidence</li> <li>• Synthesizing arguments and connecting evidence to research questions</li> <li>• Presenting results in writing, graphs or tables, etc</li> <li>• Writing a conclusion</li> <li>• Identifying limitations and challenges in study</li> <li>• Making recommendations</li> </ul>
<b>Critical-reading skills</b>	<ul style="list-style-type: none"> <li>• Developing an inquiry question</li> <li>• Extracting the central idea from a piece of text</li> <li>• Identifying purpose, audience and style of text</li> </ul>

Examples of skills to be developed	Possible formative assessments: Sub-skills or steps to build up core skills
	<ul style="list-style-type: none"> <li>• Differentiating between important and supporting information</li> <li>• Analysing arguments in text</li> <li>• Evaluating multiple sources and information</li> <li>• Comparing and contrasting the similarities and differences in the arguments and evidence in multiple sources</li> <li>• Synthesizing arguments from different sources to generate own conclusion</li> </ul>
<b>Critical-writing skills</b>	<ul style="list-style-type: none"> <li>• Deciding on a topic, title or central idea of writing</li> <li>• Determining the purpose, audience and style of text</li> <li>• Listing main arguments and counterarguments</li> <li>• Corroborating each argument with evidence and supporting information (eg examples, data, facts, etc)</li> <li>• Examining and comparing the importance and reliability of multiple sources of information</li> <li>• Citing reliable sources and listing the references</li> <li>• Organizing and structuring arguments in writing</li> <li>• Writing a conclusion</li> </ul>
<b>Technology and tools skills</b>	<ul style="list-style-type: none"> <li>• Knowing when and when not to use technology (or a tool)</li> <li>• Understanding the purpose and use of the technology or tool</li> <li>• Applying it or using it skillfully for the intended purpose</li> </ul>

The emphasis on the concurrent development of skills, concepts and content is shifting the way teachers visualize and develop their courses of study. Instead of planning the instruction and assessment in a linear progression with unit-based or content-based assessments, teachers and students are planning their courses like an organic map where learners build on and further strengthen their skills to manage dynamic, interconnected and sometimes transdisciplinary concepts, content and contexts.



## An example of a skills-based approach in DP business management

This section provides a range of examples of how to use a skills-based approach in learning and assessment in a DP subject. The aim is to support students to develop and build up the primary and ancillary skills needed to reach their learning objectives in a course of study and beyond.

### Anticipating the possible inferences in assessment design

In effective assessment design, one needs to have a clear goal of “what to assess” and to anticipate the possible inferences during the design and planning stage of the assessment activity, and not just at the marking or feedback stage. A skills-based approach could support teachers and learners to make these professional judgements on “what to assess” and make more fine-tuned inferences that support effective feedback on individual strengths and needs.

For example, student A and student B may obtain the same grade in an extended written task or practice question in business management. In a skills-based approach to assessment, breaking down the extended task into varied assessment activities may help student A understand that they have a relatively good understanding of business tools and theories and could make use of these in analyses, while they have yet to improve their skills to write business recommendations and for citation and referencing. While student B may be very strong in their quantitative and graphing skills, they seem to encounter more challenges in their reading comprehension, especially when reading multiple texts on business cases within a limited amount of time.

Planning the formative assessment activities with the purpose of making inferences would help to design effective assessments that generate helpful formative feedback to learners and teachers. It could also help to enhance student agency by empowering students to engage in peer review and self-assessments.

### A fluid and responsive approach to formative assessment

The examples in the “Unpacking the primary and ancillary skills” section below are by no means exhaustive and should not be followed in a prescriptive way. What makes a skills-based approach to learning and assessment fascinating is the fluidity and responsiveness of the approach and how the course of study is organized like an organic map, with limitless possibilities in how IB teachers and students can plan and shape the course in response to learning and teaching needs. The teachers and learners exercise their agency to see what skills are required at different times of the course of study, and how to build in these skills seamlessly as students engage with various concepts, content and contexts. The role of the teacher is then to tease out what skills students have mastered and those that may require further practice, and to plan opportunities for formative assessments where students can scaffold these skills and engage in constructive feedback with one another.

Using principles of “task, observations and conversations” in formative assessments, teachers sometimes isolate one or two of these skills in formative assessment, and design varied tasks for learners with different strengths and levels of progress. This flexible and responsive approach to assessment is one of the reasons why formative assessment is more like a creative process and dialogue instead of a rigid manual to follow. This is also why a skills-based approach to assessment can fit well with the emphasis on student and teacher agency that is central to the IB’s education philosophy.

# Unpacking the primary and ancillary skills

## Examples of primary skills in DP business management

### What are primary skills?

- Primary skills are usually stated explicitly in the learning objectives of the course of study and assessment objectives of the IB summative assessments.
- Primary skills are subject-specific, demonstrating the key concepts and content in the subject.
- In IB summative assessments, the IB tries to minimize assessment biases and remove barriers to learning by designing assessments that assess these key constructs only and reduce the ancillary skills that are assessed.
- In formative assessments, IB teachers design a variety of tasks to help them make inferences about student understanding and they use these to inform the next steps in learning and teaching. These may include:
  - designing assessments that focus on individual, specific skills
  - providing concrete and specific feedback to students
  - multimodal assessments
  - using conversations, observations and tasks
  - anticipating and identifying common misconceptions or challenges and supporting students to remove these barriers to learning
  - promoting student agency and collaboration
  - tracking progress in learning.

Examples of primary skills (What inferences can be made about students' understanding?)	Examples of tasks, conversations and observations to assess these skills
<ul style="list-style-type: none"> <li>• Defining and analysing business terms, tools and theories</li> <li>• Applying business terms, tools and theories in case studies</li> <li>• Quantitative skills                             <ul style="list-style-type: none"> <li>Forecasting numbers</li> <li>Understanding data and graphs, including infographics</li> <li>Calculating business-related data, eg income and expenses</li> </ul> </li> <li>• Preparing and constructing a statement of accounts</li> <li>• Explaining benefits and costs, processes or methods in business cases</li> <li>• Evaluating advantages and disadvantages of one or more aspect of a business</li> <li>• Understanding business concepts (eg ethics, creativity, sustainability and change) and applying them in analyses when it is helpful</li> <li>• Making business recommendations based on sound rationale or evaluation</li> </ul>	<ul style="list-style-type: none"> <li>• Draw a concept map or graphic organizer (eg Frayer model to assess conceptual understanding)</li> <li>• Short-response questions on business terms and concepts</li> <li>• Student-designed quizzes</li> <li>• Draft an outline</li> <li>• Short written tasks</li> <li>• Fill in the blanks in a statement of account or evaluate a sample statement of account</li> <li>• Show and tell</li> <li>• Extracts from newspapers and short analysis</li> <li>• One-minute business pitch</li> <li>• Presentations</li> <li>• Discussions</li> <li>• Debates or role-play exercises on making business decisions or recommendations</li> <li>• Evaluate a sample business proposal written by a peer and suggest ways to improve it</li> </ul>

Examples of primary skills (What inferences can be made about students' understanding?)	Examples of tasks, conversations and observations to assess these skills
<ul style="list-style-type: none"> <li>• Social entrepreneurship skills</li> <li>• Evaluating and critically analysing different aspects of a social enterprise business</li> <li>• Making business recommendations                             <ul style="list-style-type: none"> <li>• Critically evaluating the situation (eg challenges and opportunities) faced by an enterprise</li> <li>• Proposing one or multiple solutions to a seen challenge</li> <li>• Connecting recommendations to the evaluation</li> </ul> </li> <li>• Writing and organization                             <ul style="list-style-type: none"> <li>• Presenting business analysis and evaluation in a structured manner in writing</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Comparative judgment: Rank sample business proposals and explain why one is better than the other</li> </ul>

### What are ancillary skills?

- Ancillary skills are implicit skills that are often transdisciplinary and transferrable to other subjects.
- Ancillary skills are not usually the key learning objectives and assessment objectives in the subject, but they are necessary skills for the students to perform effectively in the subject and in authentic assessment tasks.
- Many ancillary skills, such as research and citation skills, are important for students to succeed in the subject and beyond, and support their metacognition development as lifelong learners.
- In IB high-stakes assessments, the IB tries to minimize assessment biases by reducing these ancillary skills (for example by reducing the amount of reading material in exams).
- The goal of formative assessments is to remove barriers to learning so that **all** students can access the curriculum and summative assessments fairly, and to promote their skills as lifelong learners. It is therefore sometimes helpful to convert these ancillary skills to primary skills and design formative assessments where they become the key constructs.
- Some students may require explicit training on some ancillary skills, while others do not. It may also be helpful to refer students to further resources on ancillary skills for them to engage in self-directed learning. Formative assessments may help students and teachers make these decisions.
- In formative assessment, IB teachers design a variety of tasks to help them draw inferences about student understanding and to support learning. The tasks follow the same principles as those listed for primary skills. These tasks may include:
  - designing assessments that focus on individual, specific skills
  - providing concrete and specific feedback to students
  - multimodal assessments
  - using conversations, observations and tasks
  - anticipating and identifying common misconceptions or challenges and supporting students to remove these barriers to learning
  - promoting student agency and collaboration
  - tracking progress in learning.
- Effective assessments are those that are domain-specific, therefore engaging students in subject-specific texts and content (for example, students' reading comprehension skills will improve through engaging with business management articles and vocabulary).

<b>Examples of ancillary skills</b> (What inferences can be made about students' understanding?)	<b>Examples of tasks, conversations and observations (engaging with business management texts and content)</b>
<b>Reading comprehension</b> <ul style="list-style-type: none"> <li>• <b>Ability to comprehend multiple sources within a limited time (reading and processing speed)</b></li> <li>• <b>Working memory: Recalling different facts while processing the text and the case study</b></li> <li>• <b>Understanding vocabulary that students may not be familiar with</b></li> <li>• <b>Understanding main ideas from text</b></li> </ul>	<ul style="list-style-type: none"> <li>• Observations: How long did it take each student to read the texts?</li> <li>• Draw an outline or timeline of events to summarize texts</li> <li>• Conversation: Questions on comprehension</li> <li>• Student-generated reading questions</li> </ul>
<b>Summarizing and synthesizing skills</b> <ul style="list-style-type: none"> <li>• <b>Synthesizing data and main arguments from multiple sources</b></li> <li>• <b>Presenting main ideas in a concise manner</b></li> <li>• <b>Filtering out irrelevant or less important information</b></li> </ul>	<ul style="list-style-type: none"> <li>• One-paragraph summary</li> <li>• One-minute presentation</li> <li>• Draft an outline (five bullet points)</li> </ul>
<b>Writing and organization</b> <ul style="list-style-type: none"> <li>• <b>Writing extended responses</b></li> <li>• <b>Supporting arguments with evidence or examples</b></li> <li>• <b>Presenting ideas in a structured manner in writing</b></li> </ul>	<ul style="list-style-type: none"> <li>• Write a short essay on an assigned topic</li> <li>• Peer-assessment or evaluate a sample student essay and suggest ways to improve</li> </ul>
<b>Graphing skills</b> <ul style="list-style-type: none"> <li>• <b>Drawing graphs and charts</b></li> <li>• <b>Manipulating data in graphs and charts</b></li> </ul>	<ul style="list-style-type: none"> <li>• Observations: How long does it take each student to draw a graph or chart using available data? Where do they seem to get stuck?</li> <li>• Drawing a graph or chart using available data</li> <li>• Manipulating data in graph using spreadsheets</li> </ul>
<b>Mathematical skills</b> <ul style="list-style-type: none"> <li>• <b>Converting text into tables and mathematical calculations</b></li> <li>• <b>Understanding numbers and data from a table</b></li> <li>• <b>Arithmetic and understanding percentages</b></li> <li>• <b>Knowing when and how to use a calculator or not</b></li> </ul>	<ul style="list-style-type: none"> <li>• Draw tables from texts</li> <li>• Evaluate a sample table or dataset drawn by a peer and suggest ways to improve it</li> <li>• Linkage to formative assessments in DP mathematics</li> <li>• Yes-No-Maybe-Why? activity (when to use a calculator)</li> </ul>
<b>Citation and referencing skills</b> <ul style="list-style-type: none"> <li>• <b>Differentiating between relevant and irrelevant information in texts</b></li> </ul>	<ul style="list-style-type: none"> <li>• Conversations, Yes-No-Why? questions to differentiate between relevant and irrelevant data</li> </ul>

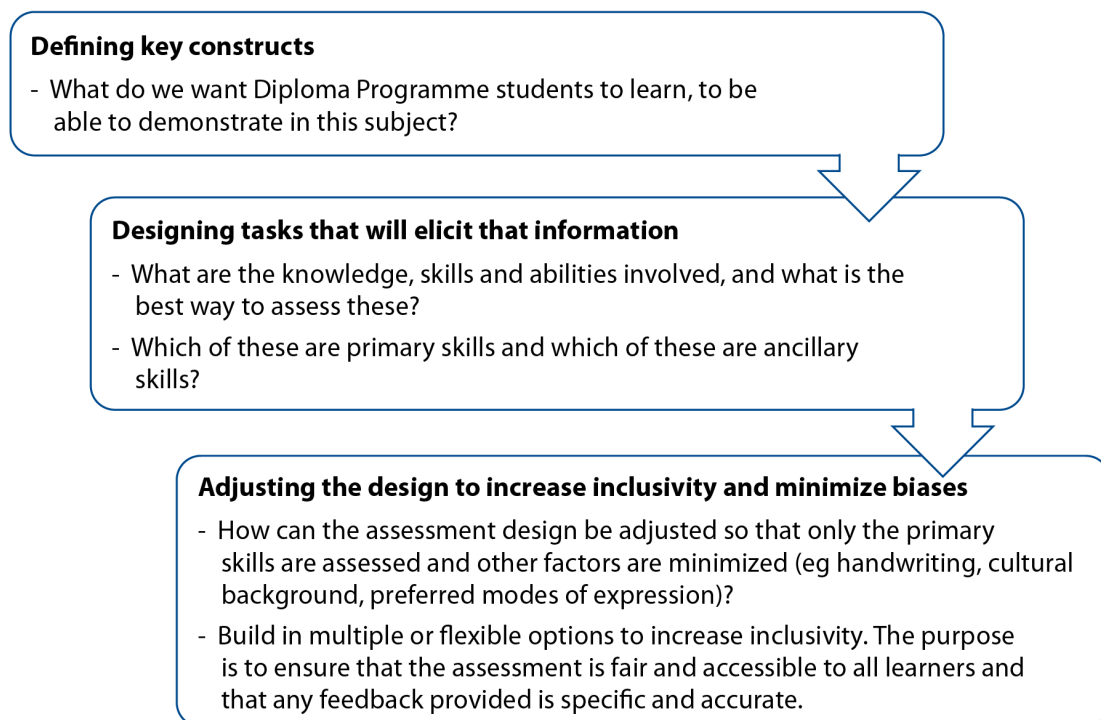
<b>Examples of ancillary skills</b> (What inferences can be made about students' understanding?)	<b>Examples of tasks, conversations and observations (engaging with business management texts and content)</b>
<ul style="list-style-type: none"> <li>• <b>Using examples, facts and figures in texts to support arguments</b></li> <li>• <b>Implicit use and explicit citation of sources</b></li> <li>• <b>Citing sources and data and synthesizing with own knowledge</b></li> <li>• <b>Knowing what is and is not there in the sources, and how to add own perspectives when citing sources</b></li> </ul>	<ul style="list-style-type: none"> <li>• Matching exercises (matching arguments with evidence from texts)</li> <li>• Writing sentences to cite from texts</li> <li>• Highlighting texts in colours and categorizing information by themes or arguments</li> <li>• Collaborative wiki project with reference list</li> </ul>
<p><b>Research skills</b></p> <ul style="list-style-type: none"> <li>• <b>Searching and selecting relevant information on the internet</b></li> <li>• <b>Linking evidence to what has been learned in the course of study (topics and themes)</b></li> <li>• <b>Generating case studies to support each other's learning (student-led assessments, peer-assessment and feedback)</b></li> <li>• <b>Ability to collaborate with peers and produce individual work</b></li> </ul>	<ul style="list-style-type: none"> <li>• Create two folders on a file-sharing platform, one with relevant sources from the internet, another with examples of interesting but irrelevant sources</li> <li>• Create an annotated bibliography</li> <li>• Highlight relevant points in sources</li> <li>• Draw a concept map to link examples to key concepts and themes</li> <li>• Student-led assessments (peer- and self-assessment on case studies)</li> </ul>

## Universally-designed assessments

Another principle central to assessment and measurement knowledge is to ensure that assessments are fair to all learners and minimize unintended biases in assessment tasks. Universal design of assessments may support teachers to make these judgments and to design fair and inclusive assessments for all learners (figure 19). For further details, please refer to the National Center on Educational Outcomes ([https://nceo.info/Assessments/universal\\_design/overview](https://nceo.info/Assessments/universal_design/overview)).

Figure 19

Key steps in the universal design of assessment



An assessment task “is biased if it gives an advantage to one person without that being the point of the assessment” (*Assessment principles and practices—Quality assessments in a digital age* 2019: 44). This bias could arise during the marking or feedback process or could be related to the design of the assessment question itself.

Teachers might have an unconscious bias when evaluating student work, such as inadvertently preferring student work with neater handwriting, or they may be affected by other factors, such as the gender, nationality or the cultural background of their students (*Assessment principles and practices—Quality assessments in a digital age* 2019). Moreover, teachers’ grading of a student work is often biased by the first impression they have of that piece of work or with their prior experience with the student, known in research as the halo effect.

For example, Malouff et al (2014) found that the prior experience with a student’s oral presentation significantly biased the subsequent grading of the written work completed by that student. The halo effect may be even stronger in assessments for formative purposes, where teachers have plenty of prior

experience with students and their work, and where assessment may be biased by additional factors such as whether the student was perceived to be motivated, attentive or making an effort in class, possibly leading to a cumulative bias.

What is the concern with bias in formative assessment if the results are not “high stakes”? One reason is that bias in formative assessment not only leads to assessment tasks and outcomes that are “unfair” but, more importantly, it affects the quality of feedback, preventing students and teachers from making the best use of feedback in guiding future learning and teaching.

Bias can be positive as well as negative. An assessment that has a positive bias may provide a student or a group of students with an unintended advantage. While a student may welcome the better results or more positive feedback caused by a positive bias, the bias inevitably limits the potential of the assessment to reflect what the student knows and does not know, therefore missing the opportunity to help students identify areas to develop further and plan next steps in learning. This was illustrated in an example shared in an interview with a DP language and literature teacher.

“I’ve done a systematic analysis of the school assessment results and discrepancies between predicted versus actual grades in my subject over the last few years. ... Students who are native speakers tend to be overmarked or over-predicted than the students who are not responding in their first languages. I wonder if this reflects a positive bias towards native speakers in my class, and perhaps some of the formative assessments we have used reflected their fluency in English instead of the actual language and literature skills that we want to target? This is something I want to explore more in future designs of formative assessments.”

In addition to bias that occurs during the evaluation of a formative task, bias could also be related to the assessment question itself. This could be connected to the context in which the question was set, favouring students with a certain kind of exposure or from a certain language or sociocultural background. It could also be related to the format and type of assessment tasks. For instance, essay questions or project work may unfairly advantage a certain group of students, while multiple-choice or short-response questions may favour another (*Assessment principles and practices—Quality assessments in a digital age* 2019).

Biases can be subtle and not easily identifiable in an assessment task. One way DP teachers could try to minimize assessment bias is to identify if there are certain sub-groups in class that tend to underperform in the subject or in certain types of assessment activities. The teachers can then tease out the primary and ancillary skills that the assessment questions evaluate to see if any of those are unrelated to the test construct, and if the question could be designed in a different way to minimize bias. For further details, refer to the “[Unpacking the primary and ancillary skills](#)” section.

In a class with students from diverse backgrounds and with different strengths, interests, backgrounds and learning needs, it is not always possible to remove all potential biases from formative assessment. Yet what teachers can do is to adopt a “balanced approach to assessment design, using a variety of different types of assessment task and format” to ensure students with diverse strengths and needs have equal opportunities to demonstrate their understanding (*Assessment principles and practices—Quality assessments in a digital age* 2019: 46).

This balanced approach aligns with the principles of universal design for learning by encouraging multiple means of engagement, representation, action and expression (Hall et al 2012). Teachers can design questions and stimulus materials in a wide range of formats, and also design the tasks to enable students to respond and express what they know in multiple modalities.

Finally, an important aspect of minimizing bias is to ensure that students with learning support requirements are not discriminated against unfairly in formative assessment. For example, when a teacher uses essay questions predominantly to assess formatively students in the first year of the DP, this may produce a disadvantage to students with dyslexia or who are responding in their second or third language. The presentation of a large quantity of multimedia texts or stimulus materials in a timed assessment may also have an unintended bias against students with slow processing speed or those with poor executive function, such as students with attention deficit disorder. The use of figures of speech or humour in stimulus materials of non-literature subjects may present an unintended disadvantage towards students who are on the autistic spectrum or who have theory-of-mind impairment.

While the results of formative assessment are not “high stakes”, the presence of bias in these tasks, especially when cumulative in nature, could be demotivating for these students, preventing them from



gaining the most from feedback, and potentially hampering future learning and teaching. It is important to ensure DP teachers have the essential assessment literacy to minimize bias in formative assessment and to provide appropriate allowances or access arrangements to those students when necessary (*Assessment principles and practices—Quality assessments in a digital age* 2019).

## An example of minimizing biases in formative assessment: Looking at primary and ancillary skills

A DP business management teacher has been using some practice questions to elicit evidence on the abilities of her DP year 1 students to evaluate different aspects of a business, and to apply business terms, tools and theories in their evaluation. Students were asked to read a number of short texts in class and to write an extended response to the question. The teacher found that the group of second language learners in class tended to underperform in this type of question. She also noticed that a few of the students who did not have prior training in academic writing seemed to find the task more challenging. The teacher then teased out the primary and ancillary skills assessed in these tasks.

<b>Primary skills</b>	<ul style="list-style-type: none"> <li>Evaluating advantages and disadvantages of one or more aspect of a business</li> <li>Applying business terms, tools or theories in evaluation</li> </ul>
<b>Ancillary skills</b>	<b>Reading comprehension</b> <ul style="list-style-type: none"> <li>Ability to comprehend multiple sources within a limited time (reading and processing speed)</li> <li>Working memory: Recalling different names, roles and events while processing the text and the case study</li> <li>Understanding vocabulary that may be unfamiliar</li> <li>Understanding main ideas from text</li> </ul>
	<b>Writing and organization</b> <ul style="list-style-type: none"> <li>Writing extended responses</li> <li>Supporting arguments with evidence or examples</li> <li>Presenting ideas in a structured manner in writing</li> </ul>
	<b>Citation and referencing skills</b> <ul style="list-style-type: none"> <li>Using examples, facts and figures in texts to support arguments</li> <li>Implicit use and explicit citation of sources</li> <li>Citing sources and data and synthesizing with own knowledge</li> <li>Knowing what is and is not there in the sources, and how to add own perspectives when citing sources</li> </ul>

The teacher wondered if some of the students underperformed in this task because they had a disadvantage in the ancillary skills and not necessarily in their subject knowledge and skills. She decided to make the task fairer by reducing the reading load and by providing students with a visual aid that included the key details of the business case to reduce the load on the students' working memory. The teacher also built in choices to the formative assessment by allowing students to submit their responses in the format of an extended written task, an infographic or recorded as a podcast. This helped the teacher have a better understanding about whether a student was underperforming because of an ancillary skill or if there was truly a gap in understanding and evaluating the business cases.

The teacher also designed a formative assessment activity to understand further her students' strengths and the gaps in their referencing and citation skills. The students were asked to provide feedback on a sample student essay and then to engage in self-assessment on their own written assignments, using a set of criteria that focused mainly on referencing and citation skills. The teacher and students were able to identify their strengths and where they needed more help. For example, it was noticed that some students



would like further training on how to add their own perspectives when citing sources. For further details, refer to the [“Unpacking the primary and ancillary skills”](#) section.

## Student agency in assessments: Peer- and self-assessments

Is formative assessment just another word for good pedagogy, the art and science of teaching? This touches on a fundamental question driving the educational philosophy in the IB: **Who leads and evaluates formative assessments?** As the IB highlights the importance of the student-centred classroom, there is a common understanding in the IB community that formative assessment goes beyond pedagogy: it is not something that is only designed and assessed by the teacher but is an interactive and dynamic process that is led and co-owned by both learners and teachers.

The idea of student agency plays a central role in formative assessment. In the conceptual framework first put forward by Black and Wiliam (2004) and later elaborated by Wiliam and Thompson (2017), peer-assessment and self-assessment form two of the five key strategies in formative assessment.

	Where the learner is going	Where the learner is right now	How to get there
<b>Teacher</b>	Clarifying and understanding learning intentions and criteria for success	Eliciting evidence of student understanding through effective class conversations, observations and learning tasks	Providing feedback to both learners and teachers to move learning forward
<b>Peer</b>		Peer-assessment and feedback: Activating students as instructional resources for one another	
<b>Learner</b>		Self-assessment: Activating students as owners of their own learning	

Adapted from Wiliam and Thompson (2017)

Formative assessment is collaborative in nature and promotes peer learning and feedback and collaboration with teachers (Topping 2009; Srijbos, Wichmann 2018). For IB students to benefit from peer- and self-assessments, teachers also need to support students to internalize success criteria and enhance their assessment capabilities or “evaluative judgment” (Carless, Boud 2018; Panadero et al 2019). In other words, students will be better prepared to understand “what is good” and to reflect on the quality of their work and that of their peers.

“Given the criteria, they mark it like teachers, they mark their fellow students' work, and then give an appraisal of the work. But you'll always give them a structure, always touch the positive before the critical element of it. So, the students would say, I like your research question, I like how you analyse the case and the conclusion is balanced, but you really missed out on evaluating it more, because the way you do is very one-sided [...] I found that they learn so well as peers.”

—IB educator elaborating on the process of peer-assessments

Student-centred formative assessment shapes the way teachers design and deliver assessment and it also changes the language teachers use. Instead of seeing formative assessment as something driven and led by teachers, it is now understood as assessment where learners and teachers have equal voices and agency. This approach may change fundamentally how some teachers and students approach learning and assessment in a more traditional classroom. The learner now assumes a much more active role in the design and evaluation of assessment tasks, such as co-constructing success criteria, giving feedback and sometimes presenting counterpoints to the teachers' evaluation of their work. The learners also take up an important role in evaluating the effectiveness of formative assessment strategies by providing feedback to their teachers. The ability of learners to share qualitative feedback and engage in deep conversations with their peers and teachers enables them to internalize and act on the feedback.

## Insights from research on peer-assessments

- Teachers can devise a specific assessment rubric to target the strengths and gaps in understanding observed in the task. Students then use the assessment rubric to provide qualitative feedback to one another (Chuck, Young 2004).
- Allow time for students to discuss the assessment criteria, to adapt it for their own use or to co-construct the success criteria (Prins et al 2005; Topping 2009).
- Students need practice to be able to engage in meaningful peer-assessments and feedback.
- Sometimes it may be helpful to allow two peer assessors to assess one student's work separately. This allows students and the teachers to see if the feedback is aligned (Topping, 2009).
- Calibrated peer review has been reported to improve the reliability of peer-assessments (Carlson, Berry, 2003).
  - In the standardization process, each student rates a few assignments that were already rated by the teacher.
  - This standardization process allows students to align their standards with one another and with the teacher through engaging in calibration and qualitative feedback.
  - The standardization process helps to build a common understanding in class of "what good looks like".
- Provide opportunities for students to act on the feedback immediately and to revise and resubmit their work (Chuck, Young 2004).
- Some students find it easier to engage in peer-assessment anonymously to provide honest and helpful feedback to one another without worrying about personal relations. One way to do this is through digital platforms where students can submit their work anonymously.
- Students need explicit support on how to give constructive and actionable feedback but also on how to receive and cope with feedback (Prins et al 2005).
- Students benefit from instructional scaffolds on how to engage in peer feedback, such as:
  - identifying the strengths and gaps in understanding
  - justifying comments
  - providing suggestions for improvements (Strijbos, Wichmann 2018; Tsivitanidou et al 2018).
- The class can talk explicitly about how they can improve the process of peer-assessment next time (Topping, 2009).

## Insights from research on student self-assessments

- The formative purpose of self-assessment needs to be made clear: it is usually to generate helpful feedback to improve drafts or work in progress. It is not used for a summative purpose to determine one's own grade (Andrade, Valtcheva 2009).
- Criteria-related self-assessment is a promising practice of self-assessment (Andrade, Valtcheva, 2009).
  - In the process, students gather information about their own performance or learning process, and compare it explicitly to success criteria, goals or standards.
  - Students have the chance to revise and resubmit their work.
- Self-assessment practices are highly effective in improving student performance when integrated with training in self-regulated learning (Hudesman et al 2013).
  - Students understand that their performances on tasks are not only determined by their level of subject knowledge and skills but also by their learning strategies and how they choose to approach a problem.
  - In self-assessment, students focus on questions they have answered incorrectly and reflect on both the academic content and specific strategies used.

Students suggest alternative strategies to approach the problem and try to solve the original problems again.

- Other useful techniques include asking students to reflect on their level of confidence in a task or in a question. They can then compare this with the qualitative feedback or marks obtained to see if their original perception aligned with that of their teachers and peers.

## Metacognition and self-regulated learning in formative assessments

A key role of teachers in formative assessment is to support learners to develop their assessment capabilities. The goal is for students to acquire the experience, skills and self-efficacy needed to become active agents in formative assessment. Supporting students to develop their metacognition and self-regulated learning through peer- and self-assessment is an important way to achieve this (Clark 2012; Panadero et al 2018).

Quality feedback encourages learners to make their tacit knowledge explicit, such as their intentions, ideas, opinions, beliefs and skills. Students benefit most from formative feedback that guides them on how to perform a task more effectively, rather than praise, rewards or feedback that simply ask them to “try harder” or “redo their responses” (Hattie, Timperley 2007). This is why feedback that focuses on students’ metacognition is particularly helpful to generate formative insights.

### Making metacognition explicit in formative assessments

- Many teachers and education researchers are experienced with using taxonomies in assessment to support the development of success criteria. Many commonly used taxonomies focus on cognitive skills, such as the SOLO taxonomy or the revised Bloom’s taxonomy, which the IB command terms are based on.
- Marzano and Kendall’s new taxonomy includes metacognitive skills and self-system thinking in addition to four levels of cognitive skills (figure 20). This adds a helpful dimension when designing tasks, observations and conversations for formative purposes. For example, how do we measure and track progress in students’ abilities to:
  - set goals for themselves
  - monitor their learning processes
  - monitor the clarity and accuracy of their answers or responses
  - examine their own motivation and efficacy in a task
  - examine their emotional response to feedback or possible setbacks in learning
  - evaluate why it is important to learn a certain skill or topic?

Figure 20

Marzano and Kendall's new taxonomy

<b>Self-system thinking</b>		Examining importance Examining efficacy Examining emotional response Examining motivation
<b>Metacognition</b>		Specifying goals Process monitoring Monitoring clarity Monitoring accuracy
<b>Cognitive skills</b>	<b>4. Knowledge utilization</b>	Investigating Experimenting Problem-solving Decision-making
	<b>3. Analysis</b>	Specifying Generalizing Analysing errors Classifying Matching
	<b>2. Comprehension</b>	Symbolizing Integrating
	<b>1. Retrieval</b>	Executing Recalling Recognizing

## Student-generated questions and assessments

Another way to promote student assessment capability is with the use of student-generated questions or assessments, where students take ownership of their learning by designing the assessment task or questions themselves. The students then decide how to evaluate these tasks and co-construct the criteria and provide feedback to their classmates through peer-assessments. In some IB World Schools, student-led assessments are an important part of formative assessments in DP classrooms (Choi 2020).

Students require instructional scaffolds to develop assessment skills that are developmentally appropriate. For example, students need to think deliberately about questions such as the following.

- What or which assessment objectives are we assessing?
- How do we assess the assessment objective? How do we write a good question that is an appropriate challenge for my peers?
- How do we define “good” in this question? What do we want to see in a quality piece of student work?
- What success criteria will we use?
  - Is there a markscheme or correct answer to some of these questions?
  - What about the open-ended questions? What do we expect to see in a quality response?
- What are some of the possible misconceptions or barriers to learning that our class may have on this topic, concept or skill? How do we discern between student responses that:
  - demonstrate full understanding of the concept
  - make a slip or careless mistake
  - have a gap in understanding
  - have a significant misconception (Bennett 2011)?

These are complex questions for students to answer, but with more practice, guidance and feedback from teachers, students can gradually become more skilful in thinking about the questions. We cannot expect students to demonstrate assessment capabilities at the level of their teachers, but engaging students in these student-led assessments and making judgments about assessment is itself a learning experience and what is sometimes called “assessment as learning”. One of the purposes of engaging students in student-generated assessment is to train their evaluative judgment, their ability to understand “what is good” and to reflect on the quality of their work and that of their peers.

## Promising practices in IB schools: Examples of student-led assessment

### Self-assessment and reflection

Subject	Example	IB World School
<b>DP chemistry</b>	<i>Mid-term exam—Gap analysis</i>	Branksome Hall, Canada
<b>DP language A: language and literature</b>	Self-assessment on essay writing (Paper 2)	Toronto French School, Canada
<b>DP physics</b>	Self-assessment on group discussion ( <i>Roundtable on Bohr's atomic spectra</i> ) Student sample	Toronto French School, Canada

### Peer-assessment

Subject	Example	IB World School
<b>DP language A: Language and literature</b>	Success criteria and form to guide peer assessment and feedback	Toronto French School, Canada

### Student-generated questions and assessments

Subject	Example	IB World School
<b>DP mathematics: Applications and Interpretation</b>	Student-designed quizzes and peer feedback	Toronto French School, Canada
<b>DP physics</b>	Student-designed assessments: Understanding assessment objectives and writing IB-style questions	Toronto French School, Canada



## Formative assessments and the ATL skills

Developing assessment capabilities in learners is strongly related to the ATL in the IB. Effective formative assessments could be used to support the development of thinking skills, research skills, communication skills, social skills and self-management skills in IB learners. At the same time, IB teachers often use formative assessments to elicit evidence of student growth in these ATL skills and to track progress.

How can ATL skills be assessed? First, it is important to look at what each of these qualities mean and unpack them in terms of sub-skills (see table below). Literature has also indicated that interpersonal and intrapersonal skills are often “content and context bound”, hence the assessment of these skills should be made “within a meaningful context” (Kießling et al 2015). In other words, it is more valid and construct-relevant to measure students’ communication and social skills in an authentic context where learners need to interact and communicate with one another, sometimes known as “performance tasks”, instead of relying on indirect measures such as a traditional paper and pencil test or self-reported questionnaires. Therefore, many IB teachers adopt formative assessment strategies, such as direct observations, peer-assessments and reflection tasks, to assess ATL skills in the IB classroom.

The following table lists some examples where ATL skills could be integrated in formative assessments in IB classrooms.

ATL skills	Examples of peer-assessment or self-assessment questions <i>In this task/activity/group discussion, am I able to/are my peers able to ...?</i>
<b>Thinking skills</b>	<ul style="list-style-type: none"> <li>• formulate a reasoned argument to support opinions or conclusions</li> <li>• propose a new perspective, solution or approach to an issue</li> <li>• show understanding of multiple and unfamiliar viewpoints when formulating arguments</li> <li>• make use of knowledge from a different subject</li> <li>• make linkages to TOK?</li> </ul>
<b>Research skills</b>	<ul style="list-style-type: none"> <li>• formulate a research question that is focused and open for inquiry</li> <li>• validate the quality and reliability of sources and compare contradictory information</li> <li>• differentiate between relevant and irrelevant information and prioritize what is important to respond to the research question</li> <li>• understand the importance of academic honesty and acknowledge use of sources through citations and references</li> <li>• know what is and is not there in the sources, and how to add own perspectives when citing sources</li> <li>• link evidence to what has been learned in course (topics and themes)?</li> </ul>
<b>Communication skills</b>	<ul style="list-style-type: none"> <li>• listen attentively and respectfully to others</li> <li>• explain viewpoints or ideas verbally or in writing</li> <li>• adapt the way of speaking and writing, such as use of different vocabulary and examples, when conversing with different audiences</li> <li>• check for understanding of the other person and adjust ways of expression in case of confusion or misunderstanding</li> <li>• check for own understanding and seek for clarification or elaboration</li> </ul>

ATL skills	Examples of peer-assessment or self-assessment questions <i>In this task/activity/group discussion, am I able to/are my peers able to ...?</i>
	<ul style="list-style-type: none"> <li>give and receive helpful feedback?</li> </ul>
<b>Social skills</b>	<ul style="list-style-type: none"> <li>consider alternative points of views or take the perspectives of others</li> <li>discuss multiple ideas or perspectives and come up with a shared understanding</li> <li>resolve a conflict or misunderstanding</li> <li>reflect on the impact of our words and behaviours on others</li> <li>make informed choices and decisions in social situations and support others to do the same?</li> </ul>
<b>Self-management skills</b>	<ul style="list-style-type: none"> <li>manage time, set priorities and meet deadlines</li> <li>set own learning goals</li> <li>organize tasks and break down a larger task into specific steps</li> <li>revise and improve on drafts of work</li> <li>maintain motivation and look for personal relevance in tasks</li> <li>discuss and use learning strategies that could best support our learning</li> <li>learn from mistakes and stay resilient in face of challenges</li> <li>receive feedback and use it to improve learning?</li> </ul>

Besides peer- and self-assessment, there are various approaches to assess ATL skills. Depending on the purpose of assessment, IB teachers exercise their professional judgments to determine what is the best method to assess these competencies in different contexts. Some of these approaches, such as those to assess interpersonal skills, have been well documented in literature. These include self-report measures, validated rating scales, standardized situational judgment tests, online portfolio assessment, scenario-based assessment and performance-based assessment such as role-play activities (National Research Council 2011; Rider, Keefer, 2006; Scheffer et al 2008). These assessment approaches were reported to differ in their degree of validity, reliability, construct relevance and fairness. As in any assessment activity, one important consideration is to balance all these principles to derive a fit-for-purpose assessment. The following table shows some examples in which IB teachers designed creative and meaningful assessments to assess formatively ATL skills in their subjects.

## Examples of effective practices: Assessing ATL skills

Subject	Example	ATL skills	IB World School
DP chemistry	<i>Harkness discussion on the nature of science: Assessing communication skills and application</i> <sup>1</sup>	Communication skills	Branksome Hall, Canada
DP mathematics: Analysis and approaches	<i>Group experiment: Assessing communication and application skills through conversations and observations</i>	Research skills Communication skills Social skills	Toronto French School, Canada
DP physics	<i>Self-assessment: Reflecting on processing skills, creative/critical-thinking skills and communication skills in roundtable discussion</i>	Research skills Thinking skills Communication skills	Toronto French School, Canada

Subject	Example	ATL skills	IB World School
	<i>Teacher observations of a science roundtable: Assessing communication and application skills</i>		
Any	<i>Harkness discussion rubric: Assessing group discussions</i> <sup>2</sup>	Thinking skills Communication skills Social skills	

Notes 1, 2: The Harkness Discussion was originally developed at Phillips Exeter Academy as an approach to learning based on discussions and student-centred learning and teaching.

## Formative assessments and the IB learner profile

As well as ATL skills, another aspect that is closely related to formative assessment is the IB learner profile. While not always assessed explicitly in formative assessment, quality assessments and feedback support the development of the learner profile attributes. The following table shows some questions related to the IB learner profile that teachers may consider when designing and delivering quality formative assessments.

<p><b>Inquirers</b></p> <ul style="list-style-type: none"> <li>• How can assessments be designed to encourage curiosity in students and support them to approach problems or questions in creative ways?</li> <li>• How can assessments be designed and feedback be shared to motivate students' love of learning?</li> <li>• Is it possible to assess students' questions instead of answers? For example, assessing student-generated reading questions or inquiry questions.</li> </ul>
<p><b>Knowledgeable</b></p> <ul style="list-style-type: none"> <li>• Are there opportunities for students to engage with issues and ideas that have local and global significance?</li> <li>• How can the development of conceptual understanding be assessed?</li> <li>• Is there a good balance between concepts, content and contexts to assess students' application and synthesis of knowledge and skills?</li> </ul>
<p><b>Thinkers</b></p> <ul style="list-style-type: none"> <li>• How can assessments be designed that encourage creative, critical and complex thinking?</li> <li>• How can meaningful and challenging formative assessments be designed to assess higher-order reasoning skills, instead of the ability to memorize and comprehend facts?</li> </ul>
<p><b>Communicators</b></p> <ul style="list-style-type: none"> <li>• Should assessment opportunities be provided for students to demonstrate and improve their communication and social skills?</li> <li>• How can student ability to collaborate, respect and listen to the perspectives of others be assessed?</li> </ul>
<p><b>Principled</b></p> <ul style="list-style-type: none"> <li>• Are current approaches to formative assessment fair to all students?</li> <li>• How can assessments be designed so that students respect fairness and demonstrate academic integrity in the process?</li> </ul>
<p><b>Open-minded</b></p> <ul style="list-style-type: none"> <li>• Are current formative assessments designed and represented in ways to reflect multiple perspectives?</li> <li>• How can formative assessments be designed and evaluated so that they reward students' abilities to seek and evaluate a range of perspectives?</li> </ul>
<p><b>Caring</b></p> <ul style="list-style-type: none"> <li>• Can formative assessments be designed to encourage and value students' abilities to show empathy, compassion and respect?</li> </ul>
<p><b>Risk-takers</b></p> <ul style="list-style-type: none"> <li>• Do current formative assessments and feedback encourage students to take challenges in their learning, and to approach uncertainty with forethought and determination?</li> </ul>

- Is a positive learning and assessment environment present, where students feel it is okay to make mistakes and stay resourceful and resilient in face of challenges?
- How can formative assessments be designed so that they encourage students to explore new ideas and innovative strategies?

### **Balanced**

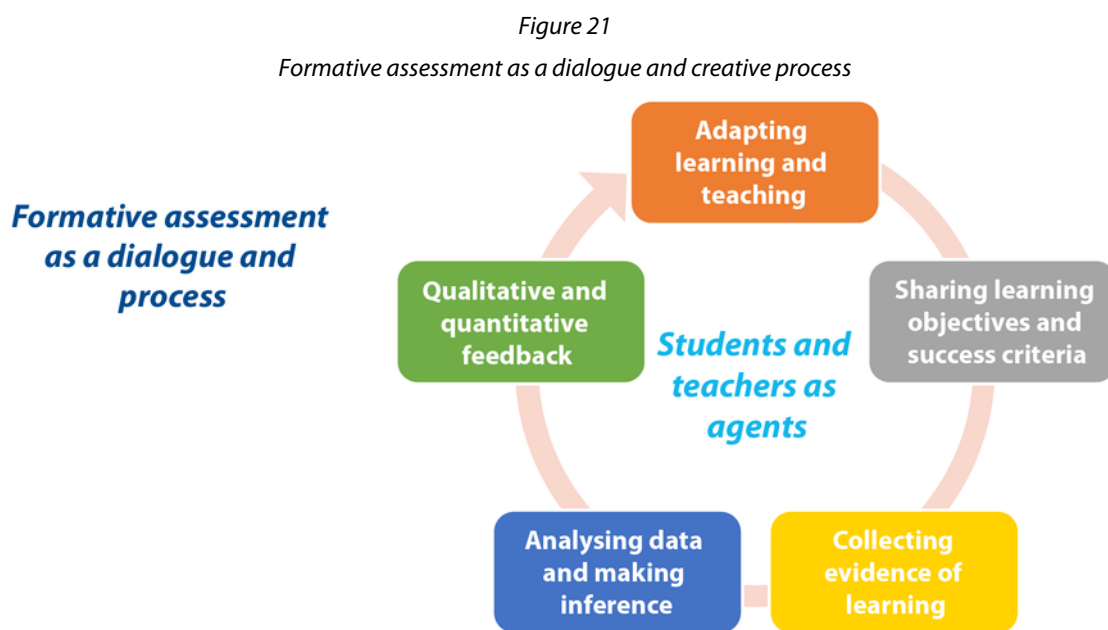
- Do current formative assessments and feedback reflect the value of a balanced and well-rounded development in students, including their social and emotional growth?

### **Reflective**

- How can students' abilities be assessed to reflect on their own thinking and experience?
- What makes a good piece of reflection? How can students be encouraged to be thoughtful about their strengths and weaknesses and to use this to guide their learning?

## How can digital technology be used to enhance and transform formative assessments?

How is digital technology transforming assessments in IB classrooms? Does digital technology enhance the way students and teachers share learning objectives, collect evidence of learning, analyse data and provide personalized feedback and feedforward (figure 21)? Could digital technology potentially redefine formative assessments by allowing teachers to assess knowledge and skills that were previously hard to assess in a paper and pencil task?



Digital formative assessment has become a more prominent part of learning and teaching in IB classrooms, perhaps even more so as many schools have adopted some form of remote learning and teaching during the COVID-19 global pandemic. In the pilot study, we visited DP classrooms and talked with many DP educators and students and came across a wide range of digital assessment practices. DP teachers and students are exploring creative and effective ways to engage in tasks, conversations and observations on digital or online platforms, such as:

- using online quizzes and short-response questions on game-based apps
- using learning platforms to share success criteria and mark student assignments
- engaging students in peer- and self-assessments
- using online blogs and shared documents so that students can interact with one another and engage in spontaneous feedback with their peers and teachers.

## An example of a digital formative assessment

### **Subject: Spanish B**

In the language acquisition class, not only do students learn about digital texts such as blogs as a new genre of text, they also engage with these digital texts directly in formative assessments and peer feedback.

In this class, students were asked to write a blog post related to the cultural identity of an island located in Peru, attending to the text type, organization of paragraphs, cohesive devices and the positive and negative aspects related to the effects of tourism and the theme “identities” (figure 22).

The teacher explained that the use of an interactive blog format allowed the students and her, to provide instant feedback and recommendations to each other, and to make revisions at the same time. This helped the teacher assess elements that are hard to analyse with a paper and pencil task, such as students’ abilities to interact with one another and to respond to feedback immediately.

Figure 22

Interactive blog in a Spanish B course

The screenshot shows a digital workspace for a Spanish B course. At the top, it says "Y-12 Spanish B- Identidad cultural" and "Made with a quick smile". The workspace is divided into several columns:

- Emeli:** Contains a "TASK" section with instructions like "Checking content and grammar of a blog related to Cultural Identity" and "Text type- Blog".
- Blog de:** Features a title "Visita el lugar de sus sueños, un lago con magia: Titicaca" and a paragraph describing the lake as a magical place. A green highlight says "No olvides mencionar cuando fuiste a ese lugar."
- Blog:** Has a title "Titicaca: Es más valioso que tu dinero." and a paragraph. A green highlight says "EN LA INTRODUCCION: No olvides mencionar el país, y conectores para cambiar de idea."
- Blog:** Has a title "Titulo - Explorando la rica cultura en las islas de Titicaca" and a paragraph. A green highlight says "No olvides mencionar cuando fuiste a ese lugar."
- Blog: Mi experiencia iluminada en el lago titicaca:** Has a title and a paragraph. A green highlight says "Cuando fuiste a ese lugar?".
- ¡El Misterio de Titicaca, Explicado!:** Contains text about the best and worst aspects of the lake.
- Aspectos positivos- negativos?:** A section for discussing positive and negative aspects.

Example contributed by Emeli Ward, NIST International School, Thailand

In education, the framework of replacement, amplification and transformation (RAT) is sometimes used to describe the impact of technological integration on curriculum, instruction and learning (Hughes et al 2006; IBO 2015). To date, there is limited published research or a conceptual framework that elaborates on the impact of digital technology on educational assessment. This section builds on the RAT framework to highlight how digital technology may enhance or transform the way teachers assess and what we assess in DP subjects.

Figure 23

Impact of digital technology integration on educational assessment (RAT framework)

How digital learning and assessment may enhance or transform:			
assessment of the abilities and skills			what is assessed
Type of response	Type of stimulus	Assessment task design	Test construct and success criteria
<p><b>REPLACEMENT</b></p> <ul style="list-style-type: none"> <li>• Typing instead of writing</li> <li>• Drag and drop</li> <li>• Assistive technology such as speech to text</li> <li>• Drawing concept or mind maps, and drawing connections between stimulus</li> <li>• Highlight in text instead of writing</li> </ul>	<ul style="list-style-type: none"> <li>• Assistive technology such as screen reader or digital dictionary to reduce access barrier</li> <li>• Reducing amount of text or reading load in non-language subjects</li> </ul>	<ul style="list-style-type: none"> <li>• Ability to analyse or evaluate multiple multimedia sources at the same time</li> <li>• Digital portfolios and learning journals (eg ability to reflect on learning process and strategies)</li> <li>• Assessing referencing skills and ensuring academic integrity (eg software to detect plagiarism)</li> <li>• Peer assessment and feedback using comparative judgment</li> </ul>	<ul style="list-style-type: none"> <li>• Intuitive problem-solving</li> <li>• Ability to work with novel problems and open-ended projects</li> <li>• Collaboration in problem-solving or inquiry project</li> <li>• Other higher-order thinking skills such as decision-making and asking good questions</li> <li>• Digital literacy (eg finding relevant information on the internet; differentiating between what is important and useful and what is not)</li> <li>• Metacognition and learner attributes (eg ability to learn from mistakes and persist)</li> </ul>



As demonstrated in the examples in figure 23, the impact of digital technology integration on how **students are assessed** can be further analysed in three aspects: the type of responses collected, the type of stimulus presented, and the design of the assessment task itself. The digitization of assessment may enhance the type of responses collected and stimulus presented in assessments, allowing multiple representation of knowledge and skills. These changes usually occur at a lower degree of transformation, such as **replacing** a function that already exists in paper and pencil tests. For example, these changes allow students to type or highlight in text instead of writing, reducing the amount of text or reading load in a question. Changes in the stimulus and response types may also **amplify** the current functions of an assessment—for example, assessing students' abilities to analyse and synthesize various multimedia sources simultaneously. Digital technology may also transform the assessment task design, changing the way the assessment is delivered, adapted or how feedback is shared. An example of this would be the use of adaptive testing and automatically generated feedback, or the use of digital learning platforms to personalize learning and assessment for individual students.

The integration of digital technology could also potentially transform **what is assessed**, thereby redefining the test constructs and influencing the definition of “good” or “success” in these assessments. Some examples of **transformation** would be when teachers make use of a simulation online to assess students' abilities to understand complex ecosystems and make predictions, or the use of virtual reality to evaluate students' ability to interact in authentic everyday conversations in their second language.

While transformation in assessment could occur with the use of advanced technologies, it can also be achieved with much simpler and accessible forms of digital technology that are already present in many IB classrooms. Some examples might be:

- using shared documents or digital portfolios to assess collaborative research and writing skills
- using the review functions in shared documents to evaluate students' progress and assess their ability to respond to feedback and improve their work in multiple drafts.

What is transformative in these tasks is not how advanced or elaborate the technological functions are, but that the digital approach could potentially allow assessment of elements that were previously reported as difficult to assess via paper and pencil methods, especially in an examination context. Examples would be interactive conversational skills, process and collaboration skills, systems thinking, and projecting trends. Overall, although changes at the response and stimulus levels may occur alone without involving changes at a deeper level, a transformation at the test construct level almost always involves a redesign of the assessment task, and thus the type of stimulus presented and responses collected.

Due to the nature of formative assessments, there is huge potential in the formative space for teachers to design and try out digital assessments that are **transformative** in nature. This may allow for the assessment of skills and knowledge that were considered difficult to assess in a paper-based examination.

## Supporting IB teachers to become assessment-capable

The earlier sections in this guide and teacher support material have outlined various assessment principles, approaches, strategies and examples of promising practices for formative assessments. There is no right way or single approach to formative assessments, and the effectiveness of these strategies or practices depends on the contexts and the unique backgrounds, strengths, needs and interests of the students. Assessment is about making professional judgments on what to assess and how to assess. It is also about making decisions on how to draw inferences on learning outcomes and what feedback to share to move learning forward. Therefore, instead of holding on rigidly to what may be the most effective or promising practices for formative assessment, it is more important to support and empower teachers and learners to make these complex judgments on their own. The IB's vision on formative assessment is that all learners and educators can exercise their assessment capabilities and agency to make the best decisions about assessment in their contexts.

Various education assessment organizations and research studies have attempted to define what is teacher assessment capability or assessment literacy (DeLuca et al 2019; DeLuca, Klinger 2010; Michigan Assessment Consortium 2020; Xu, Brown 2016). In this section, it is broadly defined as the teachers' "capacity to negotiate assessment knowledge and skills in the context of their classroom teaching" (DeLuca et al 2019: 159). In this publication, various research and evidence-based practices and promising examples from IB World Schools have been shared. Together, they contribute to what the IB understands to be assessment capabilities in teachers and learners. Assessment capabilities in IB teachers and learners incorporate questions on the following:

- What is the purpose of assessment, and how can high-quality assessments that are fit for purpose be developed or selected?
- What are the key concepts or skills to assess, and what is the best way to assess those targeted skills or concepts?
- What are the inferences that can be drawn from assessments, and what level of confidence is there about the assessment outcomes?
- What makes feedback effective, how is the effective feedback and feedforward shared to move learning forward?
- How can assessment principles be balanced—such as issues of construct relevance, fairness, reliability, comparability and manageability—to design meaningful assessments?
- How can assessment biases be identified and minimized? How can barriers to learning be removed to make formative assessments fair to all learners?
- How is student agency and the development of evaluative judgment in students supported using quality peer- and self-assessments?
- How can digital technology be used to facilitate or enhance formative assessments?
- How is a learning environment put in place and supported to enable quality formative assessments?
- How can assessment data be used to improve formative assessment practices?

## School policies and practices that support quality formative assessments

Besides supporting teachers to become assessment-capable, IB World Schools create effective culture and school policies to support quality formative assessments in their contexts. The following table shows some items in the *Programme standards and practices* that support effective assessments in schools, both for formative and summative purposes.

Standards	Practices
<p><b>Culture through policy implementation (0301)</b> Schools develop, implement, communicate and review effective policies that help to create a school culture in which IB philosophy can thrive.</p>	<p>The school implements, communicates and regularly reviews an assessment policy or policies to help create a culture of continuous learning and growth. (0301-05)</p>
<p><b>Approaches to assessment (0404)</b> Learning, teaching, and assessment effectively inform and influence one another.</p>	<p>Students and teachers use feedback to improve learning, teaching and assessment. (0404-01)</p>
	<p>The school uses assessment methods that are varied and fit-for-purpose for the curriculum and stated learning outcomes and objectives. (0404-02)</p>
	<p>The school administrates assessment consistently, fairly, inclusively and transparently. (0404-03)</p>
	<p>Students take opportunities to consolidate their learning through assessment. (0404-04)</p>

## Examples of promising practices in IB World Schools

IB World Schools operate in rich and diverse educational and cultural contexts across the world. Teachers and schools use the *Programme standards and practices* to support the development of a formative assessment policy and effective practices that best fit their contexts and vision on formative assessments. The following are some of the examples of promising practices and school policies shared by IB World Schools in the pilot study.

- Collaboration, planning and scheduling**

  - Building in time for collaboration and planning of formative assessments among teachers: vertical—across years and MYP and DP—and horizontal, across subjects)
  - Building in time for deeper learning and formative assessment in class
- Collective teacher efficacy**

  - Providing professional development activities for teachers to engage deeper with topics on formative assessment
  - Teachers as learners: Building professional learning communities to share examples and practices on formative assessments

A school culture that encourages teachers to experiment and innovate in formative assessment practices

Conducting action research to examine the effectiveness of school assessment policies and classroom assessment practices

Autonomy for teachers in designing school-based assessments. For example, a school has moved away from a traditional written test model on mid-term assessments towards a model that uses common standards, while encouraging multiple ways of assessment and representations of knowledge

- **Feedback, grading and reporting practices**

- Comments without a grade

- Qualitative feedback: Strengths and areas for improvement

- Setting the next targets for learning

- Feedback in multistage assignments

- Decoupling feedback and grades

- Reporting on the learning process and progress besides achievement

- Student agency in report cards: Students reflecting on their learning approaches and strategies and setting personal goals for learning

- **Assessment for growth**

- Assessment and feedback that promote resilience in the face of challenges

- Feedback that focuses on learning processes and strategies instead of achievement

- Quality formative assessment and feedback as guidance for future learning and growth, instead of assessing innate abilities or predicting future grades

- **Driven by data**

- Formative use of summative assessment data

- Analysing trends and patterns in assessment data

- Making assessment fair for all learners: Identifying and minimizing potential assessment biases and barriers to learning

# Glossary

Term	Definition
Academic honesty/ integrity	A set of values and skills that promote personal integrity and good practice in learning, teaching and assessment.
Analytic markscheme	A markscheme which tells you what the right answer is and where marks should be awarded.
Assessment	The collection of evidence of learning in order to make judgments about learning and teaching.
Assessment access requirements	A student with assessment access requirements is one who requires changes in assessment conditions to demonstrate his or her level of attainment.
Assessment as a dialogue	An interactive assessment process embedded in learning and teaching. It includes the cycle of sharing learning intentions, eliciting and interpreting evidence of learning, providing feedback, and using that to inform the next steps of learning and teaching. Both learners and teachers take ownership in this process.
Assessment as Learning	An approach to assessment that uses the assessment process and feedback to enhance the metacognitive skills and evaluative judgment of students: namely the ability of students to understand “what is good” and to reflect on the quality of their work and that of their peers.
Assessment bias	Assessment bias is where a defined group (for example, a racial or ethnic group or gender) performs differently on a specific question or assessment task than average for a reason other than ability in the key construct being assessed.
Assessment capability	The ability of a teacher or student to negotiate assessment knowledge and skills in the context of their classroom learning and teaching.
Assessment component	An assessment component is made up of one or more tasks that are collected together, comprising part of the overall assessment. For example, a research assignment, portfolio of work, a project, or an examination paper.
Assessment criteria	Criteria against which a student’s performance is evaluated.
Assessment for Learning	An approach to assessment that focuses on seeking and interpreting evidence, then generating feedback. This feedback is then used to improve students’ performance. Assessment for Learning is also known as formative assessment.
Assessment of Learning	An approach to assessment that is carried out at the end of a course or unit of study to provide evidence of student achievement, such as to determine the level of performance by allocating a letter or numerical grade. Assessment of Learning is also known as summative assessment.
Assessment Objective	One of a set of statements describing the skills, knowledge and understanding that will be assessed.
Assessment response	A term used to describe all material produced by a student in response to an assessment question or activity.

Term	Definition
Assessment strategy	The method or approach that is used when gathering information about student learning: for example, observations, conversations, open-ended tasks and selected responses tasks.
Assessment task	The activity or series of activities students engage in so that teachers can elicit evidence of learning and engage students in feedback on their learning.
Assessment tool	A method of eliciting evidence of learning, to gather information about a learner's performance and understanding.
Authentic Assessment	Assessments that reflect ways in which students may encounter these activities in the real world.
Backwash effect	In the educational context, backwash usually refers to the way teaching and learning is changed by how the student is assessed at the end of the course.
Command terms (IB-specific)	The words in a question, usually in summative assessments, that explain the assessment objective being assessed.
Comparability	The degree to which a particular outcome can be considered the same as another outcome. For example, if students are given different options in a formative assessment task for reasons of inclusivity, issues related to comparability may include whether the different options are of comparable difficulty, and whether they would allow the teacher to provide equal quality feedback to all learners.
Construct relevance	The degree to which the assessment actually assesses the skills and knowledge that it intends to assess. An example of a low level of construct relevance would be evaluating a student's practical skills with a written exam.
Criterion-related assessment	An assessment process based on determining levels of achievement against previously agreed criteria. The standard is therefore fixed and is not dependent on the achievement of the entire cohort of students.
Criterion referencing	A comparison of student attainment against predefined descriptions of achievement (criteria) for grading or feedback purposes.
Discipline	A branch of learning or field of academic study; a way of ordering knowledge for the purpose of instruction (known generally for practical purposes of assessment in the MYP and DP as "subjects"). Some MYP subject groups and subjects can comprise multiple disciplines. For example, the MYP subject group "arts" includes disciplines like visual arts, drama, music, media and dance. The subject "integrated sciences" includes three disciplines: biology, chemistry and physics.
eAssessment	Assessment carried out on a computer or similar device.
Evaluative judgment of students	The extent to which students recognize and understand the reasons for quality and develop their personal capability to make complex judgments about their work. In other words, students are able to understand "what is good" and to reflect on the quality of their work and that of their peers.
Examination	Examinations are assessment that takes place under controlled conditions. They are usually timed, and comprise materials or questions unknown to the students.
External assessment	Assessment that is set and marked by the IB and not by the teacher of a student.
Feedback (and feedforward)	Qualitative and quantitative feedback provided to the learner to reflect on their current performance and to indicate the next steps needed to improve learning and teaching (the latter is also known as feedforward).

Term	Definition
Formative assessment	An approach to assessment that focuses on seeking and interpreting evidence and generating feedback, where the inferences drawn from assessment outcomes relate to the kinds of actions that would best help the student to learn. It is an ongoing assessment that is embedded in curriculum and instruction. Formative assessment is also known as Assessment for Learning.
Grade	The description of student achievement. Final grades for student work range from 1 (lowest) to 7 (highest). In IB summative assessments, the grade represents the IB's judgment on the overall qualities that the candidate has demonstrated and is consistent between years and subjects.
Grade descriptors	The articulation of the qualities expected of students to achieve each grade. A grade descriptor may be specific to a subject, specific to a subject group, or general across a whole programme. In each case, a grade descriptor should describe the same characteristics. The more specific examples only explain what these descriptions mean in a subject-specific context.
Holistic criteria or holistic appraisal	An approach to evaluating a student's work which considers the work as a single outcome, rather than looking at separate elements of it individually (for example, communication, subject knowledge, quality of argument, and so on).
Inclusive access arrangements	Changed or additional conditions during the assessment process for a student with assessment access requirements. These enable the student to demonstrate his or her level of attainment more fairly and are not intended to compensate for any lack of ability.
Interdisciplinary assessment	Combining or involving two or more branches of learning or fields of academic study within a single assessment. In the DP an interdisciplinary subject is one that meets the requirements of two subject groups through a single subject. In the MYP, interdisciplinary study can be developed both within and between/ among subject groups. MYP external interdisciplinary assessment always involves multiple subject groups.
Internal assessment	Assessment carried out by teachers in the school.
Item	Smallest unit of an assessment task or question. An item could be a whole question or parts of a question.
Manageability	The degree to which the assessment and individual tasks place a workload on the student or school. Examples of manageability include the length of the assessments, the equipment or material required to deliver the assessment, or the number of assessments required in a qualification.
Mark(s)	Criterion-related numbers awarded to reflect the quality of the student's answer to the specific question asked. In IB summative assessments, the marks are aggregated to determine the final grade.
Markscheme	Guidance for awarding criterion levels for a given piece of work.
Metacognition	The awareness of one's own knowledge, such as what one knows and does not know, and the ability to understand and regulate one's cognitive processes. It is also known as "thinking about thinking".
Multimodal assessments	An assessment task or activity that allows students to express their skills and knowledge through multiple means of representation (for example, visual, audio, written, verbal, numerical and performative).
Multiple-choice question	A question where a candidate must select the correct answer from a list of supplied possible answers.

Term	Definition
Norm referencing	Where attainment is determined by comparing (referencing) the student's performance against that of the entire population for whom the assessment is designed.
On-screen examination	A formal, timed, externally produced, media-rich examination comprising a series of tasks related to the subject and designed to be answered in a secure exam environment.
Peer assessment or peer review	A structured learning process or activity for students to critique and provide feedback to each other on their work. It supports students to develop their evaluative judgment and their skills in providing feedback to one another.
Plagiarism	The representation, intentionally or unintentionally, of the ideas, words or work of another person without proper, clear and explicit acknowledgment.
Predictability	Predictability refers to the ability of teachers and learners to anticipate what will be asked in an assessment and when. Good predictability supports students in understanding what is expected of them in the assessment. It therefore reduces construct-irrelevant responses and increases assessment validity.
Question	A task or activity used to allow a candidate to demonstrate their competence in a subject.
Reliability	The degree to which the student will receive the same outcome every time his or her work is assessed. It can refer to the reliability between examiners: for example, would two examiners or peer reviewers give the same outcome for the student? It can also refer to the reliability of a single examiner: that is, does the examiner give the same outcome every time she or he looks at the student's work?
Response language	The language in which the student answers the assessment.
Self-assessment	A structured learning process or activity for students to judge their own performance realistically and to improve their work. It supports students to reflect on their learning progress and to develop their self-regulated learning skills.
Self-regulated learning	The ability of learners to understand and control their own learning environments. Self-regulated learning includes the skills to set learning goals, select strategies that help achieve those goals, monitor progress towards the goals, and examine one's own motivation in the process.
Standard	The performance which is expected to achieve a particular score, grade or assessment outcome.
Student agency in assessment	Student agency refers to the process of activating students as owners of their own learning. In the context of formative assessment, this includes engaging students actively in the process of clarifying and understanding intentions and criteria for success, and engaging students in feedback, peer review or self-assessment.
Summative assessment	Assessment aimed at determining the competency or level of achievement of a student, generally at the end of a course of study or a unit of work.
Universal Design of Assessment	The aim of Universal Design of Assessment is that all assessments should be developed with an understanding of the range of requirements that students may have, rather than treating some students differently. This is part of IB's commitment to Universal Design for Learning (UDL), to ensure all learners can access and participate in meaningful and challenging learning opportunities.



Term	Definition
Validity	The overall term that describes whether an assessment, or the purpose for which the assessment outcomes are being used, is fit for purpose.
Validity argument	The evidence and explanation for decisions made in creating an assessment which justifies that it is fit for purpose.
Weak criterion-referencing	If student attainment is compared against predefined descriptions of achievement (criteria) and the performance of previous cohorts, this is known as weak criterion-referencing (or criterion-related referencing). This is the approach the IB takes to maintaining standards.
Working languages	The languages in which the IB communicates with its stakeholders and in which it is committed to providing a range of services for the implementation of the programmes. They are currently English, French and Spanish.

# References

- A dialogue on formative assessment in the International Baccalaureate Diploma Programme*. Internal research report, July 2019.
- Adie, L, van der Kleij, F and Cumming, J. 2018. "The Development and Application of Coding Frameworks to Explore Dialogic Feedback Interactions and Self-regulated Learning". *British Educational Research Journal*. Vol 44, number 4. Pp 704–723.
- Anderson, LW, Bloom, BS and Krathwohl, D. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York, USA. Addison Wesley Longman.
- Andrade, H and Valtcheva, A. 2009. "Promoting Learning and Achievement through Self-assessment". *Theory into Practice* Vol 48, number 1. Pp 12–19.
- Assessment principles and practices—Quality assessments in a digital age*. July 2019.
- Atkinson, SP. 2013. *Taxonomy Circles: Visualizing the Possibilities of Intended Learning Outcomes*. London, UK. BPP University.
- Baird, JA, Caro, DH and Hopfenbeck, TN. 2016. "Student Perceptions of Predictability of Examination Requirements and Relationship with Outcomes in High-stakes Tests in Ireland". *Irish Educational Studies* Vol 35, number 4. Pp 361–379.
- Belshaw, D. 2017. "Structure of Observed Learning Outcomes (SOLO) Taxonomy". Diagram shared under Creative Commons CCO 1.0 Universal Public Domain Dedication. [https://commons.wikimedia.org/wiki/File:Structure\\_of\\_Observed\\_Learning\\_Outcomes\\_\(SOLO\)\\_Taxonomy.png](https://commons.wikimedia.org/wiki/File:Structure_of_Observed_Learning_Outcomes_(SOLO)_Taxonomy.png).
- Bennett, RE. 2011. "Formative Assessment: A Critical Review". *Assessment in Education: Principles, Policy & Practice* Vol 18, number 1. Pp 5–25.
- Bennett, RE and Gitomer, DH. 2009. "Transforming K–12 Assessment: Integrating Accountability Testing, Formative Assessment and Professional Support". In C Wyatt-Smith and J Joy Cumming. Eds. *Educational Assessment in the 21st Century*. Pp 43–61. Dordrecht, Netherlands. Springer.
- Biggs, JB and Collis, KF. 2014. *Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome)*. London, UK. Academic Press.
- Black, P and Wiliam, D. 2004. "The formative purpose: Assessment must first promote learning". *Yearbook of the National Society for the Study of Education*, Vol 103, number 2, Pp 20–50.
- Black, P and Wiliam, D. 2018. "Classroom Assessment and Pedagogy". *Assessment in Education: Principles, Policy & Practice* Vol 25, number 6. Pp 551–575.
- Carless, D. 2015. *Excellence in University Assessment: Learning from Award-winning Practice*. Abingdon, UK. Routledge.
- Carless, D and Boud, D. 2018. "The Development of Student Feedback Literacy: Enabling Uptake of Feedback". *Assessment & Evaluation in Higher Education* Vol 43, number 8. Pp 1315–1325.
- Carlson, PA and Berry, FC. 5–8 November 2003. "Calibrated peer review and assessing learning outcomes". Paper presented at *33rd Frontiers in Education Conference*. Westminster, CO, USA.
- Chadwick, D. 2009. Ed. *Approaches to Building Conceptual Understandings*. Wellington, New Zealand. Ministry of Education.
- Chaim, OC and Cazarini, EW. 2017. "Peer and Self-grading: Comparisons between Analytical and Holistic Rubrics in an Engineering Classroom". *The International Journal of Engineering Education* Vol 33, number 6. Pp 1798–1802.
- Choi, W. 22 May 2020. "Bringing student choice to assessment" [Blog post]. *The IB community blog*. <https://blogs.ibo.org/blog/2020/05/22/bringing-student-choice-to-assessment/>.

- Christodoulou, D. 2017. *Making Good Progress?: The Future of Assessment for Learning*. Oxford, United Kingdom. Oxford University Press.
- Chuck, JA and Young, L. 2004. "A Cohort-driven Assessment Task for Scientific Report Writing". *Journal of Science Education and Technology* Vol 13, number 3. Pp 367–376.
- Clark, I. 2012. "Formative Assessment: Assessment is for Self-regulated Learning". *Educational Psychology Review* Vol 24, number 2. Pp 205–249.
- Crooks, TJ, Kane, MT and Cohen, AS. 1996. "Threats to the Valid Use of Assessments". *Assessment in Education: Principles, Policy & Practice* Vol 3, number 3. Pp 265–286.
- Dawson, P, Henderson, M, Mahoney, P, Phillips, M, Ryan, T, Boud, D and Molloy, E. 2019. "What Makes for Effective Feedback: Staff and Student Perspectives". *Assessment & Evaluation in Higher Education* Vol 44, number 1. Pp 25–36.
- Dede, C. 2010. "Comparing Frameworks for 21st Century Skills". In J Bellanca and R Brandt. Eds. *21st Century Skills: Rethinking How Students Learn*. Bloomington, IN, USA. Solution Tree Press.
- DeLuca, C, Coombs, A and LaPointe-McEwan, D. 2019. "Assessment Mindset: Exploring the Relationship between Teacher Mindset and Approaches to Classroom Assessment". *Studies in Educational Evaluation* Vol 61. Pp 159–169.
- DeLuca, C and Klinger, DA. 2010. "Assessment Literacy Development: Identifying Gaps in Teacher Candidates' Learning". *Assessment in Education: Principles, Policy & Practice* Vol 17, number 4. Pp 419–438.
- Effective practices in formative assessment in the DP* [Unpublished pilot study]
- Elwood, J, Hopfenbeck, T and Baird, JA. 2017. "Predictability in High-stakes Examinations: Students' Perspectives on a Perennial Assessment Dilemma". *Research Papers in Education* Vol 32, number 1. Pp 1–17.
- Fong, CJ, Patall, EA, Vasquez, AC and Stautberg, S. 2019. "A Meta-Analysis of Negative Feedback on Intrinsic Motivation". *Educational Psychology Review* Vol 31, number 1. Pp 121–162.
- Gibbs, G and Simpson, C. 2004. "Does Your Assessment Support Your Students' Learning?". *Journal of Teaching and Learning in Higher Education* Vol 1, number 1. Pp 1–30.
- Greatorex, J and Malacova, E. 2006. "Can Different Teaching Strategies or Methods of Preparing Pupils Lead to Greater Improvements from GCSE to A Level Performance?". *Research Papers in Education* Vol 21, number 3. Pp 255–294.
- Hall, TE, Meyer, A and Rose, DH. Eds. 2012. *Universal Design for Learning in the Classroom: Practical Applications*. New York, USA. Guilford Press.
- Hattie, J and Brown, GT. 2004. *Cognitive Processes in asTTle: The SOLO Taxonomy (asTTle Tech. Rep.# 43)*. Auckland, New Zealand: University of Auckland/Ministry of Education.
- Hattie, J and Timperley, H. 2007. "The Power of Feedback". *Review of Educational Research* Vol 77, number 1. Pp 81–112.
- Hudesman, J, Crosby, S, Flugman, B, Issac, S, Everson, H and Clay, DB. 2013. "Using Formative Assessment and Metacognition to Improve Student Achievement". *Journal of Developmental Education* Vol 37, number 1. P 2.
- Hughes, J, Thomas, R and Scharber, C. March 2006. "Assessing technology integration: The RAT—replacement, amplification, and transformation—framework". Paper presented at *Society for Information Technology & Teacher Education International Conference*. Pp 1616–1620. Waynesville, USA. Association for the Advancement of Computing in Education (AACE).
- Jonasson, D. 2006 "On the Role of Concepts in Learning and Instructional Design". *Educational Technology Research and Development* Vol 54, number 2. Pp 177–196.
- Kiessling, C, Essers, G, Anvik, T, Jankowska, Meneses, R, Tsimtsiou, Z, Rosenbaum, M and Silverman, J. March 2015. *General Principles for the Assessment of Communication Skills*. tEACH Assessment Subgroup. <http://komunikacjamedyczna.pl/wp-content/uploads/2013/09/Ocena-kompetencji-komunikacyjnych-w-opiece-medycznej.pdf>
- Malouff, JM, Stein, SJ, Bothma, LN, Coulter, K and Emmerton, AJ. 2014. "Preventing Halo Bias in Grading the Work of University Students". *Cogent Psychology* Vol 1, number 1.

- Marzano, RJ and Kendall, JS. Eds. 2008. *Designing and assessing educational objectives: Applying the new taxonomy*. California, USA. Corwin Press.
- Medwell, J, Wray, D, Bailey, L, Biddulph, M, Hagger-Vaughan, L, Mills, G, Oliver, M and Wake, G. September 2019. *Concept-based teaching and learning: Integration and alignment across IB programmes*. <https://www.ibo.org/contentassets/318968269ae5441d8df5ae76542817a0/revised-cbtl-final-report-march-2020-2.pdf>
- Meyer, E and Land, R. 2003 *Threshold Concepts and Troublesome Knowledge: Linkages to Ways of Thinking and Practising within the Disciplines*. Occasional Report 4, ETL Project. Edinburgh: University of Edinburgh. [www.leeds.ac.uk/educol/documents/142206.pdf](http://www.leeds.ac.uk/educol/documents/142206.pdf).
- Michigan Assessment Consortium. 2020. *Assessment Literacy Standards—A National Imperative*. [https://www.michiganassessmentconsortium.org/wp-content/uploads/MAC\\_AssessLitStds\\_2017\\_9.19.17.pdf](https://www.michiganassessmentconsortium.org/wp-content/uploads/MAC_AssessLitStds_2017_9.19.17.pdf)
- National Research Council. 2011. *Assessing 21st Century Skills: Summary of a Workshop*. Washington DC, USA. National Academies Press.
- Nicol, DJ and Macfarlane-Dick, D. 2006. "Formative Assessment and Self-regulated Learning: A Model and Seven Principles of Good Feedback Practice". *Studies in Higher Education* Vol 31, number 2. Pp 199–218.
- Panadero, E, Andrade, H and Brookhart, S. 2018. "Fusing Self-regulated Learning and Formative Assessment: A Roadmap of Where We Are, How We Got Here, and Where We Are Going". *The Australian Educational Researcher* Vol 45, number 1. Pp 13–31.
- Panadero, E, Broadbent, J, Boud, D and Lodge, JM. 2019. "Using Formative Assessment to Influence Self-and Co-regulated Learning: The Role of Evaluative Judgement". *European Journal of Psychology of Education* Vol 34, number 3. Pp 535–557.
- Piaget, J. 1952. *The Origins of Intelligence in Children*. New York, USA. International Universities Press.
- Prins, FJ, Sluijsmans, DM, Kirschner, PA and Strijbos, JW. 2005. "Formative Peer Assessment in a CSCL Environment: A Case Study". *Assessment & Evaluation in Higher Education* Vol 30, number 4. Pp 417–444.
- Programme standards and practices*. October 2018 (revised April 2020).
- Rider, EA and Keefer, CH. 2006. "Communication Skills Competencies: Definitions and a Teaching Toolbox". *Medical Education* Vol 40, number 7. Pp 624–629.
- Sadler, DR. 2009. "Indeterminacy in the Use of Preset Criteria for Assessment and Grading". *Assessment & Evaluation in Higher Education* Vol 34, number 2. Pp 159–179.
- Sadler, DR. 2010. "Beyond Feedback: Developing Student Capability in Complex Appraisal". *Assessment & Evaluation in Higher Education* Vol 35, number 5. Pp 535–550.
- Scheffer, S, Muehlinghaus, I, Froehmel, A and Ortwein, H. 2008. "Assessing Students' Communication Skills: Validation of a Global Rating". *Advances in Health Sciences Education* Vol 13, number 5. Pp 583–592.
- Shepard, LA. 2017. "Formative Assessment: Caveat Emptor". In C A Dwyer. Ed. *The Future of Assessment: Sharing Teaching and Learning*. Pp 279–303. New York, USA. Routledge.
- Strijbos, JW and Wichmann, A. 2018. "Promoting Learning by Leveraging the Collaborative Nature of Formative Peer Assessment with Instructional Scaffolds". *European Journal of Psychology of Education* Vol 33, number 1. Pp 1–9.
- Taber, KS. 2013. "Revisiting the Chemistry Triplet: Drawing upon the Nature of Chemical Knowledge and the Psychology of Learning to Inform Chemistry Education". *Chemistry Education Research and Practice* Vol 14, number 2. Pp 156–168.
- Teaching and learning with technology: A guide of basic principles*. [https://ibpublishing.ibo.org/server2/rest/app/tsm.xml?doc=g\\_0\\_iboxx\\_amo\\_1512\\_1\\_e&part=1&chapter=1](https://ibpublishing.ibo.org/server2/rest/app/tsm.xml?doc=g_0_iboxx_amo_1512_1_e&part=1&chapter=1)
- Topping, KJ. 2009. "Peer Assessment". *Theory into Practice* Vol 48, number 1. Pp 20–27.
- Tsivitanidou, OE, Constantinou, CP, Labudde, P, Rönnebeck, S and Ropohl, M. 2018. "Reciprocal Peer Assessment as a Learning Tool for Secondary School Students in Modeling-based Learning". *European Journal of Psychology of Education* Vol 33, number 1. Pp 51–73.
- Ültanır, E. 2012. "An Epistemological Glance at the Constructivist Approach: Constructivist Learning in Dewey, Piaget, and Montessori". *International Journal of Instruction* Vol 5, number 2.

- VanTassel-Baska, J and Wood, S. 2010. "The Integrated Curriculum Model (ICM)". *Learning and Individual Differences* Vol 20. Pp 345–357.
- Wiggins, GP. 1993. *Assessing Student Performance*. San Francisco. Jossey-Bass.
- Wiliam, D. 2011. *Embedded Formative Assessment*. Bloomington, USA. Solution Tree Press.
- Wiliam, D and Thompson, M. 2017. "Integrating Assessment with Learning: What Will It Take to Make It Work?". In CA Dwyer. Ed. *The Future of Assessment: Sharing Teaching and Learning*. Pp 53–82. New York. Routledge.
- Xu, Y and Brown, GT. 2016. "Teacher Assessment Literacy in Practice: A Reconceptualization". *Teaching and Teacher Education* Vol 58. Pp 149–162.
- Yeager, DS and Dweck, CS. 2012. "Mindsets that Promote Resilience: When Students Believe that Personal Characteristics Can Be Developed". *Educational Psychologist* Vol 47, number 4. Pp 302–314.