Points to be noted when using this lecture material:

This lecture material includes images etc., used by the University of Tokyo with the permission from third parties, and images, etc., provided under respective licenses. Please follow the rules determined by the respective rights-holders when using the individual images.
Copyrighted works owned by the faculty members of the University of Tokyo may only be reused for non-profit or educational purposes.
Please credit the following when using this material:

UTokyo Online Education:
UTokyo Global FFDP 2022 Gabriel Hervas

東京大学
THE UNIVERSITY OF TOKYO

UTokyo ffdp GLOBAL
Future Faculty Development Program

Video for DAY 3

Assessment, CATs, & technology

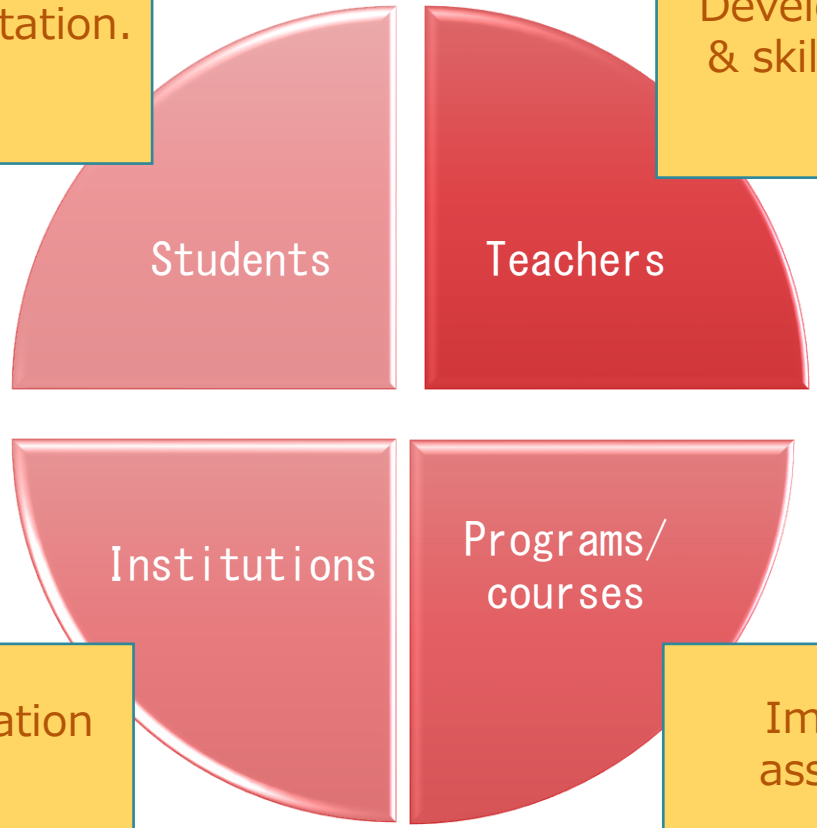
Dr. Gabriel Hervas
gabriel@he.u-tokyo.ac.jp
Center for Research and Development of Higher Education
The University of Tokyo

In this video

- Assessment vs evaluation
- Purposes
- Main questions
- Quality of assessment
- Classroom assessment techniques
- Assessing and EduTech

What can be assessed/evaluated

...we not only assess/evaluate students' learning...



Guide learning & accreditation.
Multiple agents

Develop educational knowledge & skills (SoTL) & accreditation.
Multiple agents

Accountability, accreditation quality assurance.
Multiple agents

Improving (SoTL), quality assurance & accreditation.
Multiple agents

ASSESSMENT/EVALUATION OF STUDENTS' LEARNING

Assessment and evaluation. Fundamentals

- The notions: assessment and evaluation. Perceptions.

  - Why?
  - When?
  - Who?
  - What (we determine)?
  - How?

Main questions

The notions: assessment and evaluation

Process of gathering information to…

Submit a **judgement** and determine value/quality/achievement

EVALUATION

- Grading
- Accrediting
- "Summative assessment"

Depending on the context, their use varies

Inform the learning process and make (help make) **decisions**

ASSESSMENT

- Diagnostic
- Formative
- Summative

Perceptions

Beliefs and self-beliefs impact learning (Day 1)

Also, students' **perceptions** about assessment significantly **influence** their approaches to **learning** and studying (Struyven et al., 2005).

Are teachers' and students' perceptions about assessment aligned?

Perceptions

- **Faculty** members are likely to view assessment as a **trustworthy** process **aiding** T-L.

- **Students** viewed assessment as focused primarily on **accountability** and perceived it as **irrelevant** or ignored in the T-L process.

(Fletcher et al., 2012)

The purposes of assessment

- **Diagnostic**: diagnosing departing point to adjust the T&L process.
- **Formative**: improving the learning process, at the service of learning (**low stakes**). Hand-by-hand with feedback.
- **Summative**: measuring ("final") **achievement** of learning outcomes (pass/fail; grade; etc.) (**high stakes**).

When

- **Diagnostic**: beginning of learning process (tasks, lessons, courses, etc.).
- **Formative**: across learning process. Continuous.
- **Summative**: generally, end of learning process (tasks, lessons, courses, etc.).

Who

- **Diagnostic**: mostly teacher (also, self-diagnose).
- **Formative**: teachers, peer-, and self-assessment (also other agents)
- **Summative**: mostly teacher, but we can involve peer- and self-assessment.

Assessment engagement

For all purposes and involving in decision-making about assessment

What we determine

- **Diagnostic**: previous knowledge, skills, attitudes, values, interests, etc. Learning needs at the beginning.
- **Formative**: learning progress, learning needs (across the process).
- **Summative**: learning achievement.

What we determine (another way to think about it)

- They all involve (can be thought in relation with) the assessment of learning outcomes. **Distance**.
  - **Diagnostic**: "**distance**" from the desired learning outcomes at the beginning: *do they know (do, are…) what they need to know to begin to learn and achieve the learning outcomes?*
  - **Formative**: "**distance**" from the desired learning outcomes across the learning process: *are they learning what/how they need to learn to achieve the learning outcomes?*
  - **Summative**: "**distance**" from the desired learning outcomes at the end: *have they (and at what degree) achieved the learning outcomes?*

How

Nature of content (or learning outcomes). Many techniques/instruments can be adjusted to serve diagnostic, formative, or summative purposes.

- **Conceptual**/factual knowledge: *"objective tests"/surveys, essays/open questions, concept maps, resumes, presentations, debates/forums/discussions, one-minute paper, muddiest point, …*
- **Procedures/skills/attitudes/competences**: **PRACTICE**. We assess by observation of a process or the analysis of its results/products. *Presentations, essays, laboratory work, simulation, dramatization, debates/forums/discussions, "-based" strategies, capstones, "genius-hour" works, etc.*
- **Metacognition**: *essays, self-/peer-assessment, debates/forums/discussions, reports, etc.*

*Competence: capability to apply combination of knowledge, skills, attitudes to perform successfully: discipline specific or key/soft/core/transversal (critical thinking, communication, collaboration, global citizenship, ,…).

Assessment

Comments, ideas, & doubts so far…

Take note of them, stop the video when needed.

Remember to take a break

QUALITY OF ASSESSMENT (ASSESSING ASSESSMENT)

Assessing assessment

Why?

Continuous professional development

SoTL

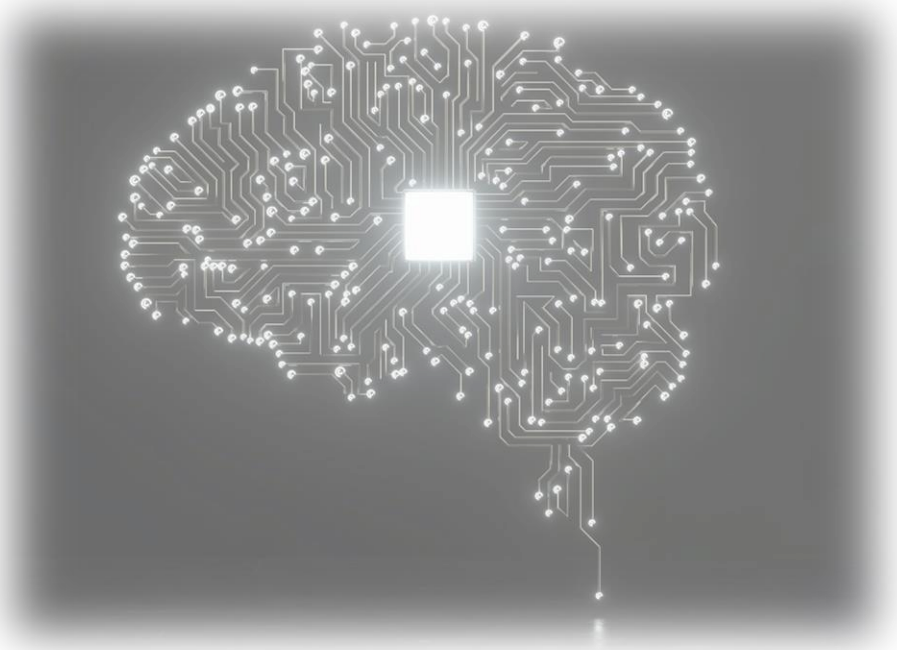
Assessing teaching quality as an attitude/duty

Contributing to students' learning

Assessing assessment

- **Fairness & flexibility**: attention to individual differences and moments of learning (equity, inclusiveness, diversity)
- **Validity**: assesses/measures what it intends to assess/measure or decisions are based on performance (credibility & transferability, from a qualitative stance)
- **Reliability**: consistent & dependable (student, rater, administration and test reliability) (dependence and confirmability, from a qualitative stance)
- **Practicality**: efficiency of design and use (time, economy, administration)

- Others:
  - **Washback**: impact of assessment over T and L
  - **Authenticity**: realistic vs real
  - **Transparency**
  - **Supportiveness of learning**



Features of quality educational assessment

- **Systematic**, comprehensive and continuous, BUT **sustainable** and feasible.
- **Participated** by students adding transparency.
- **Planned** (including the feedback), BUT adaptable and **flexible**.
- Meaningful: with **purpose** (decision making, adding value for teaching-learning).
- **Rigorous**, valid (for purpose with explicit and non-arbitrary criteria).
- **Aligned** with learning goals/outcomes and methods.
- **Cumulative**: coming back to and interleaving rather than assessing isolated blocs (retrieval!)
- Makes students' **thinking visible**.

CLASSROOM ASSESSMENT TECHNIQUES (CATS)

CATs. What are they?

During class-instruction.

Although named techniques, they are more like **short-term activities** that, in general, involve **low stakes assessment (non-summative)**

Purpose of getting information on the ongoing learning and learning process

CATS. Examples

- Tests/polls/**quizzes** before/during the class to check learning or diagnose knowledge, etc.
- Checking for understanding. **Questioning**, re-asking, rephrasing, redirecting questions, etc.
- Asking students to generate questions, **paraphrasing** ideas, retelling, summarizing, etc.
- Brief written reflections: one-sentence summaries, **one-minute paper**, muddiest point, anticipation of learning, chain notes, interactive writing, etc.
- Feedback on the **learning process**: additional questions within other assignments (difficulties, time, etc.), asking for suggestions (explicit question or anonymous box), etc.
- Making **thinking visible** in assignments: asking for additional comments arguing responses to identify gaps, etc.
- Problem/principle/**main idea recognition** tasks, steps for problem solution, etc.
- **Completing tables, pros & cons**, concept maps, knowledge/fact checklists, completing slides/sentences, etc.

CATS. Why?

- Asking students about their own learning might not be effective: "Did you understand?".

- Opportunity to **adjust** instruction/T-L to students' learning moment (although, CATs not necessarily mean an enhancement of students' learning. See Simpson-Beck, 2011).

- Opportunity for students to **think about their own learning**, seeing **T-L as a process**, & noticing our interest on their learning.

- Many of them useful when working with **large groups**, can be used **frequently** (not necessarily), can be **anonymous**, etc.

Web-references/documents with ideas

- Classroom Assessment Activities (CATs):

https://vcsa.ucsd.edu/_files/assessment/resources/50_cats.pdf
https://facultyinnovate.utexas.edu/sites/default/files/ChecksforLearning-DuringInstruction.pdf
https://teaching.berkeley.edu/resources/course-design-guide/design-effective-assessments/alternatives-traditional-testing
https://www.cmu.edu/teaching/assessment/assesslearning/CATs.html
https://cft.vanderbilt.edu/guides-sub-pages/cats/
https://teaching.utoronto.ca/teaching-support/gathering-formative-feedback/

- Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques*. Wiley.
- Barkley, E. F., & Major, C. H. (2015). *Learning assessment techniques: A handbook for college faculty*. Jossey-Bass.
- Fisher, D., & Frey, N. (2014). *Checking for understanding: Formative assessment techniques for your classroom* (2nd Ed.). ASCD.
- Simpson-Beck, V. (2011). Assessing classroom assessment techniques. *Active Learning in Higher Education, 12(*2), 125–132.

CATs…

Comments, ideas, & doubts so far…

Take note of them, stop the video when needed.

Remember to take a break

ASSESSING LEARNING & EDUC. TECHNOLOGY

How

- Learning analytics (see Tsai et al., 2018):
  - Involves collecting, measuring, analyzing information about students and their learning process to improve learning and the learning environment.
  - Data from students at all stages: performance (in assignments, class, etc.), student surveys, access to LMS, … but also from admissions process, orientations, interactions, students' characteristics/profiles, educational backgrounds, etc.
  - Data quality and management, ethical issues to access data, organizational resources/culture, etc.

How

- Automatic assessment/feedback (see Conejo et al., 2016; Crossley et al., 2016; Ochoa & Dominguez, 2020; Pinheiro Cavalcanti et al., 2021)
  - Different tools and methods to automatize assessment (formative and summative) and feedback (some are open/free). Check references.
  - Contributes to offer timely-feedback in different types of tasks (MCQ, written essays, oral presentations, etc.).
  - Time (no clear evidence that it eases teachers' workload) and resources, integration, features of quality feedback.

Educational resources/technology/tools

- WebPA (peer moderated marking system)
- Irubric, Corubrics, Google Classroom (for creating on-line rubrics)
- Skilltrack (to track and assess skills development)
- Turnitin, Speedgrader, GradeScope (evaluating, grading, feedback, etc.)
- Automatic assessment/feedback references
- Tools seen on DAY 2 (for portfolio, quizzes, etc.)

Educational resources/technology/tools

- PerUsAll (sharing comments on documents)
- Amanote (annotating course materials)
- Mentimeter, Kahoot, Socrative, Quizizz, etc. (polls, quizzes, etc.)
- Scorion, Mahara (e-portfolio)
- Peek, Goosechase (gamified field trips, missions, etc.)
- Flippity, Gimkit (creation of gamified activities)
- Mozzillla Hubs (meet, share, collaborate in 3D virtual space)
  - Virtual/augmented reality (see Radianti et al., 2020)
- VoiceThread (asynchronous voice and video talks, discussions, etc.)
- Padlet (sharing ideas, posts, threads, etc.)
- Skilltrack (to track and assess skills development)
- Turnitin, Speedgrader, GradeScope (evaluating, grading, feedback, etc.)

…

References

- Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques*. Wiley.
- Barkley, E. F., & Major, C. H. (2015). *Learning assessment techniques: A handbook for college faculty*. Jossey-Bass.
- Barton, G. & Ryan, M.E.(2014). Multimodal approaches to reflective teaching and assessment in higher education. *Higher Education Research & Development,33(*3), 409-424
- Beets, P. A. (2012). Strengthening morality and ethics in educational assessment through ubuntu in South Africa. *Educational Philosophy and Theory, 44*(sup2), 68-83.
- Biggs, J. & Tang, C. (2011). *Teaching for Quality Learning at University.* Buckingham: Open University Press/McGraw Hill.
- Birenbaum, M. & Feldman, R.A. (1998). Relationships between learning patterns and attitudes towards two assessment formats. *Educational Research, 40(*1), 90-98.
- Boud, D., Ajjawi, R., Dawson, P. & Tai, J. (eds.) (2018). *Developing evaluative judgment in Higher Education. Assessment for knowing and producing quality work*. Routledge
- Bryan, C. & Clegg, K. (Eds.) (2006). *Innovative assessment in higher education*. Routledge.
- Cano, E. & Ion, G. (eds.) (2017). *Innovative practices for Higher Education assessment and measurement*. IGI Global.
- Carless, D. (2014). Exploring learning-oriented assessment processes. *Higher Education, 69*(6), 963–976.
- Carless, D. (2020). Longitudinal perspectives on students' experiences of feedback: A need for teacher–student partnerships. *Higher Education Research & Development*, 20(3), 425-438.
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43(*8), 1315–1325.
- Carless, D., Bridges, Ka Yuk Chan, Glofcheski (Eds.) (2017). *Scaling up assessment for learning in higher education*. Springer.
- Cobeña, G. T. B., García, L. A. P., Pin, S. C. S., & Montes, L. C. Z. (2021). The formative assessment as systematic practice in higher basic education students. *International Research Journal of Management, IT and Social Sciences, 8(*2), 132-140.
- Conejo, R., Guzmán, E. & Trella, M. (2016). The SIETTE Automatic Assessment Environment. Internatio*nal Journal of Artificial Intelligence in Education, 26*, 270–292.

References

- Covington, M.V., von Hoene, L.M., & Voge, D. (2017). *Life beyond grades: Designing college courses to promote intrinsic motivation*. Cambridge University Press.
- Crossley, S. A., Kyle, K. & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods, 48*, 1227–1237
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of educational research, 83*(1), 70-120.
- Evans, C., & Waring, M. (2011a). Exploring students' perceptions of feedback in relation to cognitive styles and culture. *Research Papers in Education, 26*, 171–190.
- Evans, C., & Waring, M. (2011b). Student teacher assessment feedback preferences: The influence of cognitive styles and gender. *Learning and Individual Differences, 21*, 271–280.
- Evans, C., & Waring, M. (2011c). Enhancing feedback practice: A personal learning styles pedagogy approach. In S. Rayner & E. Cools (Eds.), *Style differences in cognition, learning, and management: Theory, research and practice* (pp. 188–203). Routledge.
- Fink, A. (2003). *The survey handbook*. Sage.
- Fisher, D., & Frey, N. (2014). *Checking for understanding: Formative assessment techniques for your classroom* (2nd Ed.). ASCD.
- Fletcher, R.B., Meyer, L.H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and Students Conceptions of Assessment in Higher Education. *Higher Education, 64*(1), 119-133.
- Harks, B., Rakoczy, K., Hattie, J., Besser, M., & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educational Psychology, 34*(3), 269-290.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81-112.
- Gierl, M. J., Bulut, O., Guo, Q., & Zhang, X. (2017). Developing, Analyzing, and Using Distractors for Multiple-Choice Tests in Education: A Comprehensive Review. Review of Educational Research, 87(6), 1082–1116.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing Paradigm in Qualitative Research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Sage.

References

- Guba, E. G., & Lincoln, Y. S. (1990). *Fourth generation evaluation* (2nd Ed.). Sage.
- Gysling, J. (2016). The historical development of educational assessment in Chile: 1810–2014. *Assessment in Education: Principles, Policy & Practice, 23(*1), 8-25.
- Ion, G., Cano-García, E., & Fernández-Ferrer, M. (2017). Enhancing self-regulated learning through using written feedback in higher education. *International Journal of Educational Research, 85*, 1–10. https://doi.org/10.1016/j.ijer.2017.06.002
- Jones, I., & Wheadon, C. (2015). Peer assessment using comparative and absolute judgement. *Studies in Educational Evaluation, 47*, 93-101.
- Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14*, 63–76.
- Knight, S., Gibson, A., & Shibani, A. (2020). Implementing learning analytics for learning impact: Taking tools to task, *The Internet and Higher Education, 45.* https://doi.org/10.1016/j.iheduc.2020.100729
- Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Sage.
- Mahabeer, P., & Pirtheepal, T. (2019). Online formative assessment tools: lecturers' experiences of using Moodle at a university in South Africa. *Journal of Educational Studies, 18*(1), 43-63
- Manzi, J., García, M. R., & Taut, S. (2021). *Validity of Educational Assessments in Chile and Latin America*. Springer.
- Meir, Eli; Wendel, Daniel; Pope, Denise S.; Hsiao, Ling; Chen, Deborah; Kim, Kerry J. (2019). Are intermediate constraint question formats useful for evaluating student thinking and promoting learning in formative assessments?. *Computers & Education, 141.*
- Newman, L. R., Roberts, D. H., & Frankl, S. E. (2019). Twelve tips for providing feedback to peers about their teaching. *Medical teacher, 41*(10), 1118-1123.
- Nilson, L. B. (2015). *Specifications grading. Restoring rigor, motivating students and saving faculty time*. Stylus.
- Oakley, B., Rogowsky, B., & Sejnowski, T. J. (2021). *Uncommon sense teaching.* TarcherPerigee.
- Ochoa, X., & Dominguez, F. (2020). Controlled evaluation of a multimodal system to improve oral presentation skills in a real learning setting. *British Journal of Educational Technology, 51(5*), 1615-1630.
- Ossenberg, C., Henderson, A., & Mitchell, M. (2019). What attributes guide best practice for effective feedback? A scoping review. *Advances in Health Sciences Education, 24(*2), 383-401.
- Pereira, D., Flores, A., & Niklasson, L. (2016). Assessment revisited: a review of research in Assessment and Evaluation in Higher Education. *Assessment & Evaluation in Higher Education, 41*(7),1008- - 1032.

References

- Pinheiro Cavalcanti, A., et al. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence, 2*. https://doi.org/10.1016/j.caeai.2021.100027
- Ponce, H. R., Mayer, R. E., Figueroa, V. A., & López, M. J. (2018). Interactive highlighting for just-in-time formative assessment during whole-class instruction: effects on vocabulary learning and reading comprehension. *Interactive Learning Environments, 26(*1), 42-60.
- Pryor, J., & Lubisi, C. (2002). Reconceptualising educational assessment in South Africa—testing times for teachers. *International Journal of Educational Development, 22*(6), 673-686.
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgennant, I. (2020). A Systematic Review of Immersive Virtual Reality Applications for Higher Education: Design Elements, Lessons Learned, and Research Agenda. *Computers & Education, 147*.
- Ramani, S., Könings, K. D., Ginsburg, S., & van der Vleuten, C. P. (2019). Twelve tips to promote a feedback culture with a growth mind-set: Swinging the feedback pendulum from recipes to relationships. *Medical teacher, 41*(6), 625-631.
- Rohrer, D., Dedrick, R. F., & Stershic, S. (2015). Interleaved practice improves mathematics learning. *Journal of Educational Psychology, 107*(3), 900-908.
- Simpson-Beck, V. (2011). Assessing classroom assessment techniques. *Active Learning in Higher Education, 12(*2), 125–132.
- Sorcinelli, M.D. (2007). Faculty Development: The Challenge Going Forward. *Peer Review, 9(*4), 4-9.
- Steen-Utheim, A., & Hopfenbeck, T. N. (2018). To do or not to do with feedback: A study of undergraduate students' engagement and use of feedback within a portfolio assessment design. *Assessment & Evaluation in Higher Education, 44*(1), 80–96.
- Stobart, G. (2010). Making a difference: Evaluating the impact of innovations in assessment. In J. Gardner, W. Harlen, L. Hayward and G. Stobart (Eds.), *Developing teacher assessment* (pp. 141-154). Open University Press/McGraw-Hill.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education, 30(*4), 325–341.
- Swan, A., Rashid, H., Meka, J., Amiel, J., & Pluta, W. (2021). Twelve tips for embedding assessment for and as learning practices in a programmatic assessment system. *Medical Teacher, 43*(3), 300-306.

References

- Tsai, Y.-S., et al. (2018). Supporting higher education to integrate learning analytics. Research report November 2018. Available at https://sheilaproject.eu/wp-content/uploads/2018/11/SHEILA-research-report.pdf
- van der Leeuw, R. M., & Slootweg, I. A. (2013). Twelve tips for making the best use of feedback. *Medical teacher, 35*(5), 348-351.
- Vásquez, A., Nussbaum, M., Sciarresi, E., Martínez, T., Barahona, C., & Strasser, K. (2017). The impact of the technology used in formative assessment: The case of spelling. *Journal of Educational Computing Research, 54*(8), 1142-1167.
- Wheadon, C., de Moira, A. P., & Christodoulou, D. (2020a). *The classification accuracy and consistency of comparative judgement of writing compared to rubric-based teacher assessment*. Retrieved at https://osf.io/preprints/socarxiv/vzus4/download
- Wheadon, C., Barmby, P., Christodoulou, D., & Henderson, B. (2020b). A comparative judgement approach to the large-scale assessment of primary writing in England. *Assessment in Education: Principles, Policy & Practice, 27*(1), 46-64.
- Winstone, N. E., & Carless, D. (2020). *Designing effective feedback processes in higher education*. Routledge.
- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*(1), 17-37.

Web-references/documents with ideas

- Classroom Assessment Activities (CATs):

https://facultyinnovate.utexas.edu/cats
https://facultyinnovate.utexas.edu/sites/default/files/ChecksforLearning-DuringInstruction.pdf
https://vcsa.ucsd.edu/_files/assessment/resources/50_cats.pdf
https://teaching.berkeley.edu/resources/course-design-guide/design-effective-assessments/alternatives-traditional-testing
https://www.cmu.edu/teaching/assessment/assesslearning/CATs.html
https://cft.vanderbilt.edu/guides-sub-pages/cats/

- Online exams:

https://academic-senate.berkeley.edu/issues/coronavirus/best-practices-remote-examinations

- Getting feedback from students

https://bokcenter.harvard.edu/getting-feedback

Thank you

Dr. Gabriel Hervas
gabriel@he.u-tokyo.ac.jp
Center for Research and Development of Higher Education
The University of Tokyo