Stimulating Reflection through Self-Assessment: Certainty-based Marking (CBM) in Online Mathematics Learning

*Sima Caspari-Sadeghi, Elena Mille, Hella Epperlein, Brigitte Forster-Heinlein
Faculty of Computer Science and Mathematics, University of Passau
Innstraße 33, 94032 Passau, Germany

Sima.caspari-sadeghi@uni-passau.de, Elena.Mille@uni-passau.de,
Hella.Epperlein@uni-passau.de, Brigitte.Forster@uni-passau.de

Abstract: *This collaborative action research highlights the need for developing students' evaluative competence and self-reflection by embedding self-and-peer assessment into online instruction. Over the course of a semester in an online master program in mathematics and computer sciences, students conducted research on assigned topics, held presentations, formulated meaningful questions for peer-assessment, and finally engaged in Certainty-based Marking (CBM) by rating how certain they are that their answer is correct. The goal of using CBM was to foster students' careful reflection and provide feedback to teachers about students' status of knowledge. A mixed-method approach was used to triangulate data from two sources: (a) assessment artifacts, i.e., student-generated questions and CBM, as evidence of learning, and (b) students' attitude captured through 'Task Perception Questionnaire'. Assessment data were analyzed by three domain experts based on their judgement of 'quality' and Kappa measure was used to assess inter-rater consistency. Quantitative analysis of questionnaire data, coupled with instructors' observation, indicated positive attitudes (engaging and useful) towards CBM among students. We conclude with a discussion of limitations as well as implications of this classroom research project.*

Keywords: Certainty-Based Marking (CBM), Online Mathematics Assessment, Self-and-Peer Assessment

1. Introduction

Empirical evidence from comparative international tests, e.g., PISA, TIMS, as well as frequent failed national educational reforms raise alarm about inadequate mathematics performance and increasing STEM disengagement, e.g., high drop-out and low enrollment rates among students. Supporting the meaningful learning of mathematical procedures and developing robust fluency with mathematical skills is an urgent priority for Western mathematics education (Foster, 2016). Teachers are supposed to play a critical role in mitigating the rift between policy, research, and practice by engaging in 'evidence-based STEM education' (Milner-Bolotin, 2018) and data-driven decision-making: to collect, analyze and use research-based data to improve education (Maxwell, 2021). This study is a classroom intervention, conducted by teacher-researcher, who used 'assessment' to simultaneously generate evidence and foster students mathematical learning.

Assessment is at the core of the learning process: it shapes how students learn and provides observable evidence of learning achievement. Traditionally, higher education focused on *'Assesment-of-Learning'(AoL):* formal, summative tests at the end of a course to measure how much students have learned. However, such once-a-year tests can not help teachers make crucial instructional decisions which need moment-to-moment information about students' progress (Stiggins, 2002). *'Assessment-for-Learning' (AfL),* on the other hand, is conducted in the classroom formatively and continuously, with the aim of supporting and improving learning through diagnosing weaknesses and problems (Wiliam, 2011). Although AfL is mostly performed by teachers, there is a call for engaging students more in assessment to become progressively independent of their teachers, e.g., Sadler (2010) urged higher education institutes to develop *'evaluative judgement'* in their graduates, the ability to judge the quality of one's own and others' work, as a sustainable life-long skill which is necessary both within and beyond higher education settings (e.g., professional jobs).

Self- and peer-assessment (SAP) is an AfL method which has the capacity to engender evaluative judgement. SAP assumes that by handing over assessment responsibility to students, they engage in active learning and become more reflective through understanding and appraising quality/standards/criteria related to work (Boud and Soler, 2016). Furthermore, interacting with criteria helps to close the gap between the current and the expected performance level. In this study, we used two SAP strategies: Student-generated Questions (SGQs) and Certainty Based Marking (CBM). By requiring students to generate meaningful, quality questions and indicate their degree of certainty (c) about the answer they choose, learners will be encouraged to reflect and self-assess their knowledge (Gardner-Medwin, 2006). This classroom study sought to answer the following questions:

Q.1. *How competent are students in producing higher-order questions for peer-assessment?*
Q.2. *How confident are students in answers they choose in self-assessment?*
Q.3. *What are the attitude and perceptions of mathematics students towards CBM?*

2. Literature Review: Certainty-based Marking (CBM)

Multiple-choice Questions (MCQs) is a widely used assessment technique which provides prompt feedback on students' learning. Although students who get the right answer might think they have knowledge and know the answer, responses to multiple-choice tests can be an evidence of knowledge as well as a pure '*lucky guess*' without any knowledge or an '*educated guess*' based on partial, uncertain knowledge. Both guesses introduce error variance into the test score and affect reliability negatively (Lindquist & Hoover, 2015). Furthermore, such chance response encourages an uncritical habit of mind in students.

To remedy this inherent problem with MCQs based on a single-best answer method, *Certainty Based Marking (CBM)*, formerly known as Confidence-based Marking, assumes knowledge is not a binary thing (you know it or don't know it), i.e., by asking 'how sure, confident, certain are you?', students start to think more carefully and look for justification and reservations. It also provides a more refined differentiation of students' knowledge levels.

Students are posed with multiple choice items. After answering, they should choose from a 3-point scale: 1 (low), 2 (mid) or 3 (high), the degree of certainty (c) about the correctness of their answers. Therefore, item score is a product of both correct answer and certainty level. Based on the reported degree of certainty, different rewards and penalties are assigned: i.e., a confident, wrong answer gets the highest penalty (see Table 1). Therefore, CBM differentiates between students who choose the same correct answer by rewarding those who can distinguish their more reliable and less reliable answers.

Table 1. Mark scheme for CBM

Degree of Certainty (c)	C=1 (low)	C=2 (mid)	C= 3 (high)
Score (correct answer)	1	2	3
Penalty (wrong answer)	0	-2	-6

Certainty-based Marking aims at (a) identification of uncertainty, (b) rewarding accurate judgement of reliability, (c) reducing biases due to over-confidence and hesitation, and (d) even diminishing unwarranted self-confidence (Gardner-Medwin, 2006).

Although CBM is used extensively in Medicine to discourage guessing in life-or-death matters (Gardner-Medwin, 2019; Nathaniel et al., 2021), several other areas also embed it in their pedagogical practices. Hassmén & Hunt, (1994) found that CBM can enhance test validity by reducing gender biases. Ehrlinger et al., (2008) studied how '*illusory over-confidence'*, in which low-ability students over-estimate their competence because they 'do not know what they do not know', could be calibrated through consistent use of CBM. In another study, Yen et al, (2010) examined the correlation between students' ability and their confidence in computer-administered MC tests. In addition to a positive association, CBM was found to be more efficient compared to traditional MC tests, because it needs fewer items to estimate test-takers' knowledge level. However, some research failed to find any positive effect on outcomes such as achievement, e.g., Foster (2021) examined the effect of repeated and formative use of CBM on summative mathematics attainment across four schools (*N=475*). A Bayesian meta-analysis of the effect sizes showed no effect on students' mathematics achievement. It was concluded that CBM cannot cause a quick, easy and visible raise in gain scores in the short time. Wu et al. (2021) suggested that CBM could be affected by individual difference variables, such as gender or risk-attitude, that are not related to the main construct (e.g., ability or knowledge).

3. Method

This action research was carried over 10 months in three phases: planning, preparation, and data collection. Although the planning phase is explained briefly, the focus of this paper will be on the 'classroom action research', as conducted by the instructors during preparation and data collection phases.

3.1. Planning phase: Collaborative Action Research

This study was conducted as a part of Faculty Professional Development Program in SKILL.de project, Germany during 2021. The goal of Evidence-based Evaluation in SKILL.de is to enhance instructors' Data Literacy: ability and competence in collecting and analyzing empirical data about students' learning in order to improve instructional decision-making. During this phase, the action research coach, a researcher in empirical learning sciences, worked collaboratively with the course instructors, a professor and her two Teaching Assistants (TA). Based on the course goals, i.e., Self-regulated Learning, they designed an action research study which embeds formative self-and-peer assessment into learning activities. A critical consideration in this phase was 'ecological validity': to make sure that intervention is a naturalistic trial, easy and low-cost to implement, without imposing any new system from outside or re-designing the whole course (Neumark, 2019).

3.2. Preparation phase: training students in assessment

This small-scale classroom research was conducted during Corona-pandemic in the online seminar "Applied Mathematics in the Math Museum", over a 14-week semester at university of Passau, Germany. The Passau Mathematics Museum encourages students to design an exhibit (i.e., applet) that communicates a mathematical concept to the visitors of the math museum in addition to delivering a scientific presentation. Course delivery was through Stud.IP (Learning Management System) as well as synchronous Zoom meetings. Participants consisted of five students in bachelor and master of mathematics and computer science. To help students become more self-regulated and control their own learning, they were asked to choose a topic from an assigned list, do research and reading on the topic, develop some questions and deliver an oral presentation. Developing students' competence to ask meaningful, quality questions and reflect deeply when answering questions are at the heart of mathematical scientific literacy. However, the results of our past study showed that students are not familiar with generating quality questions (Caspari et al, 2021).

Therefore, the first session was spent on introducing the project, getting students' consent, and instructing them about Student-generated Questions. They had no prior experience in systematically formulating questions about a topic. They were introduced to 'worked examples', a sample of questions with different quality levels (lower-order and higher-order), were encouraged to discuss what makes a good multiple-choice question (both form and functions) and were asked to judge attributes of a strong and a weak MCQ. It should be noted that not all quality features can be communicated through explicit criteria; some will remain tacit and embodied (Hudson et al. 2017). Quality levels (lower or higher) were measured with reference to Bloom's Taxonomy (1958) which stipulates different levels of cognitive complexity involved in answering the questions.

Table 2. Bloom's Taxonomy of Cognitive Levels

	Cognitive domains	Cognitive levels	Actions required
Lower-Levels	Remembering (knowledge)	Low	Recognition, recall, name, list
	Comprehension	Low	Describe, explain, summarize, visualize
	Application	Low	Use, practice, solve, manipulate
Higher-Levels	Analysis	High	Compare, deduce, analyze, infer
	Synthesis	High	Synthesize, plan, design, construct,
	Evaluation	High	Judge, criticize, estimate, justify, defend

Certainty Based Marking (CBM) was also introduced later in the course. Some studies (Bar-Hille, Budescu, & Attali, 2005) showed that students' choices of a certainty level were affected by their risk attitudes: when students have a high success probability on an item, they become risk averse (under-reporting of their certainty) and conversely become risk-taking if there is a low success probability (over-reporting of their certainty). To avoid 'demotivating' of students, we decided not to assign any score as 'penalty and reward' to certainty level. The students were asked simply to indicate their certainty level on a 3-point scale: *1(low*)= unsure/not confident; *2(mid*)= relatively confident; *3(high*)= highly confident.

3.3. Data collection phase

Students conducted self-study on a topic, prepared a presentation and formulated two MCQs which were presented at the end of their lecture. The class answered and indicated their certainty in answers (see Appendix A). There were subsequent discussions about questions (levels, ambiguity, etc.) during the whole process, instructors took some field-notes about their observation.

Students' perspectives and attitudes towards CBM were captured at the end of semester through a questionnaire, Task Perception Questionnaire (TPQ), developed by the authors. First, we reviewed existing related literature and developed an initial 7-item scale based on selective adoption of the Self-determination Theory Framework (Deci & Ryan, 1991) and Technology Acceptance Model (TAM), which are used widely to assess digital competence and acceptance (Venkatesh and Davis, 2000). The scale was reviewed by two experts (Mathematics Professor and learning science researcher) and was refined again. The final version of the TPQ is composed of five questions, on a four-point Likert scale ("strongly disagree" to "strongly agree"), measuring three aspects of a task perception, (a) *usability*: the perceived ease or difficulty in performing the task, (b) *engagement* with the task, and (c) *intention to use* in future (see Appendix B). The questionnaire was administered online and anonymously.

4. Analysis

Three mathematics instructors were instructed to code the quality of SGQ based on a two-dimensional rubric: (a) the overall quality of a question, and (b) the cognitive demand involved in a question. The overall question's quality was assessed based on its content coverage, clarity, relevance, and plausibility on a rating scale of 1-3 (1= poor, 2=good, 3= excellent). Both stems and distractors were considered. A question was rated as '*Poor=1'* if it was ambiguous, had irrelevant alternatives and very little topic coverage (Caspari et al., 2021). The cognitive demand of SGQs was measured with reference to Bloom's Taxonomy or levels of cognitive complexity (e.g., remembering; understanding; applying; analyzing; synthesizing and evaluating). The inter-rater reliability among three subject-matter experts was calculated, resulting in an overall Cohen's kappa value of *d= 0*.68 among all raters. Next, all raters and the moderator (action researcher coach) met to negotiate discrepancies. Discussions continued until consensus was reached on all codes.

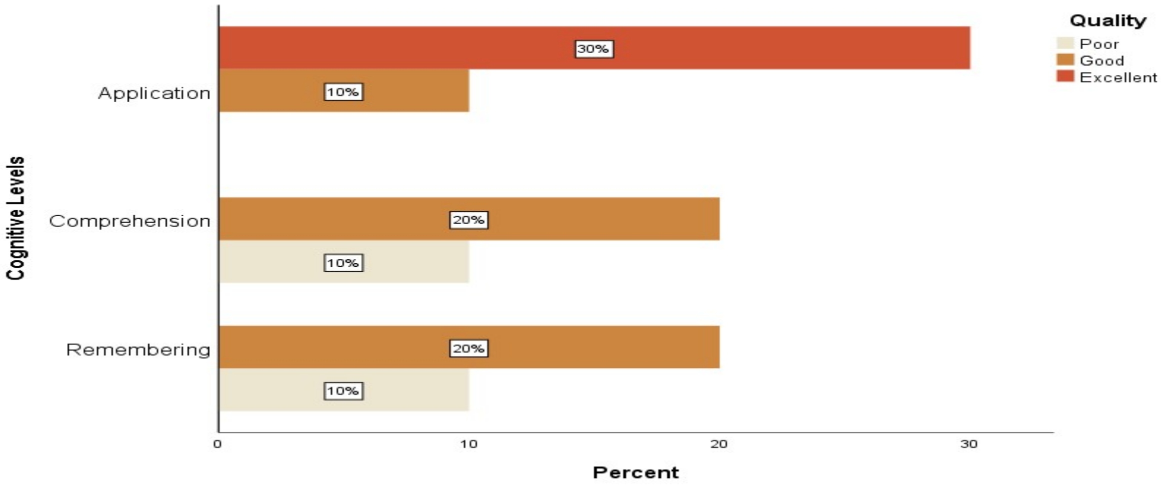
5. Results

5.1. SGQ

Q.1. How competent are students in producing higher-order questions for peer-assessment?

Results of SGQ showed all questions authored by students were at the so-called '*lower-levels'* of cognitive complexity, namely 30% at *Level 1= Remembering*, which requires mere retrieval of facts and information, 30% at *Level 2= Comprehension*, which requires understanding of materials, and eventually 40% targeting *Level 3= Application*, which necessitates the use of knowledge to perform or solve problems. None of SGQs reached '*higher-levels*' of cognitive complexity, such as analysis, synthesis, or evaluation. Quality-wise, 30 % of produced questions were rated as '*excellent*', with another 50% as '*good*' and only 20% were assessed as '*poor*'.

Figure 1. SGQ Quality



4.2. CBM

Q.2. How confident are students in answers they choose in self-assessment?

24 Out of 40 answers to all SGQs were correct. In reporting their degree of certainty in the correct answers, 30% expressed a low level of confidence, while 54% were almost/relatively sure about the correctness of their answers. Only 16% had a high degree of certainty. None of the students expressed full assurance (being 100% confident) in answers they selected (see table 3).

Table 3. Degree of Certainty in correct answers

Certainty levels	C=1 (low)	C=2 (mid)	C= 3 (high)
Assessment of correct answer	*30%*	*54%*	*16%*

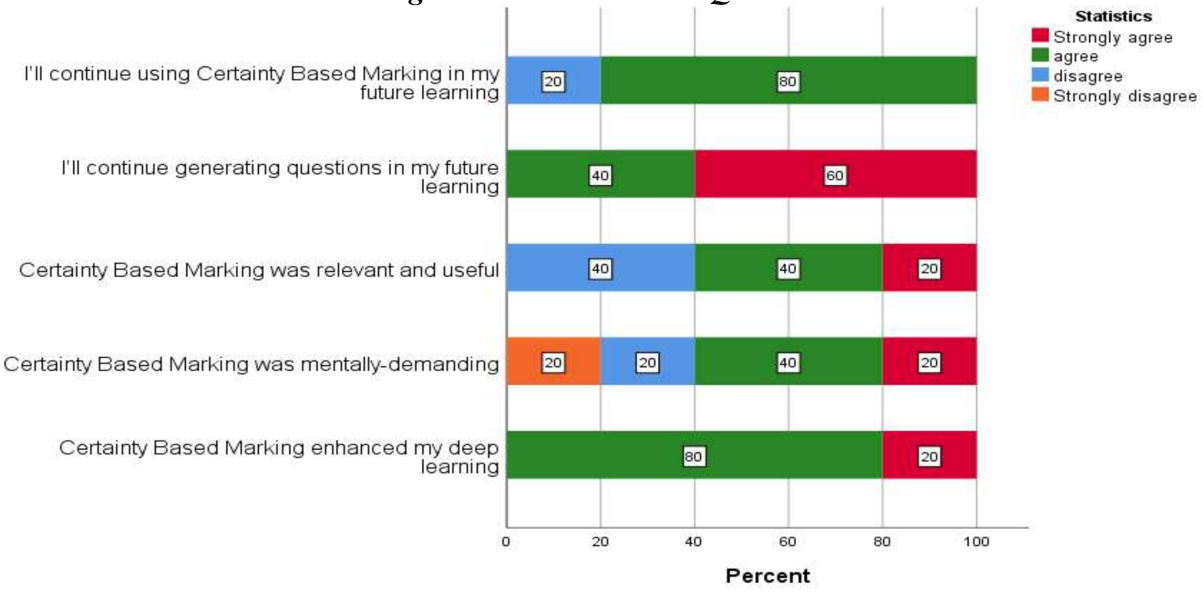
1= low confidence/unsure; 2 = average confidence/relatively confident;3= high confidence

4.3. Task Perception Questionnaire (TPQ)

Q.3. What are the attitude and perceptions of mathematics students towards CBM?

Figure 2 shows students' responses from the online anonymous Task Perception Questionnaire (TPQ). '*Usability of CBM'* received a mixed reaction: although 60% considered it as a difficult and mentally demanding task, the rest believed it was easy and manageable. Most of the students (60%) viewed the activity as relevant and useful. In terms of ‘engagement’, or to what extent CBM involved students in self-reflection, all students (80% agree, 20% strongly agree) believed CBM enhanced their deep learning and reflection. Majority of cohort expressed their positive attitude toward ‘*intention to use’*: with 80% agreed that they’ll continue using CBM in their future learning, while all students either strongly agreed (60%) or agreed (40%) about continuing the use of SGQ for future learning.

Figure 2. Task Percetion Questionnaire



5. Discussion

The analysis of data, presented in section 4, indicates that this small-scale intervention could enhance students’ participation in assessment. Survey results imply a positive attitude and students’ increased motivation to take charge of their own-and-peer assessment, a sustainable skill which is transferable to other contexts.

We found a slight improvement in the quality of SGQ compared to our previous classroom research (Caspari-Sadeghi et el, 2021). It might be tempting to ascribe this enhancement to some explicit actions taken by instructors, such as direct instruction about quality of MCQs or assigning few scores to motivate students’ serious involvement. However, due to its naturalistic design and inherent lack of control of pre-existing variables, e.g., background knowledge, action research avoids establishing any cause-and-effect relationship. Even though students failed to produce questions at higher-levels of cognitive complexity, a closer look into the existing literature and the nature of SGQ can shed some lights on this phenomenon. In a large-scale review of MCQs across the U.S. biology courses, Momsen et al., (2010) found that 90% of items are at the lowest two levels of the Bloom’s taxonomy, namely remembering and understanding. This could be partly attributed to the ‘nature’ of such questions: MCQs are often criticized for their inability to target ‘*conceptual understanding’* and being mostly focused on recall of factual knowledge (Biggs &

Tang, 2011). Additionally, developing such higher-order competences requires more time, practice, and a shift in the culture of educational systems.

Results of CBM revealed that majority of cohort (70%) were certain and sure about the correctness of their answers. It should be also mentioned that in our small sample (*N=5*), we couldn't observe students who gave incorrect answers and expressed a high certainty about their incorrect belief. Overall, our findings are in line with other studies (e.g., Sparck, Bjork, & Bjork, 2016) that suggest CBM as a useful and efficient self-test provided that it is used continuously in the classroom. There were some limitations to this study.

5.1. For reliability purposes, it would have been better to develop a longer questionnaire. For pragmatic reasons, authors decided against this, e.g., the intervention was supposed to be non-invasive and small-scale. Furthermore, the students were already assigned to several other tasks (i.e., presentation, SGQ, CBM, digital exhibits, etc.) as well as participating in a university-led survey.

5.2. This exploratory case study is more like a formative experiment carried over a short period of time. Authors make no claim over generalizability or causality of such a small-scale intervention. Cautions should be taken in attempting to replicate in other contexts.

6. Conclusion

This case study aimed to explore the development of evaluative judgement through self-and-peer assessment. Based on a classroom action research, we examined implementation, uptake as well as students' attitude towards effectiveness of CBM and SGQ as efficient techniques to engage students with assessment. It's safe to say both instructors and students believed this formative intervention effectively enhanced their learning. Although the results of the survey revealed positive attitude, we could not establish the extent to which SGQ and CBM improved students' mathematics attainment (i.e., final score). There is a need for more research on several aspects of CBM that we didn't cover in this study, e.g., Novak (2017) asserted Asian cultures find it quite unnatural to rate themselves above the average. It might be interesting to examine if other demographic variables such as 'discipline' or 'socio-economic class' might have any implications for using CBM.

Acknowledgement: Funding was provided by German's Federal Ministry of Education and Research (BMBF).

References

[1] Bar-Hillel, M., Budescu, D., & Attali, Y. (2005). Scoring and keying multiple choice tests: A case study in irrationality. *Mind & Society*, 4, 3-12.

[2] Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University*. Maidenhead, UK: Open University Press.

[3] Bloom, B.S., Engelhart, M.B., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational goals* (Handbook 1: Cognitive domain). New York: Longmans Green.

[4] Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400-413.

[5] Caspari-Sadeghi, S., Forster-Heinlein, B., Mägdefrau, J., and Bachl, L. (2021). Student-generated Questions (SGQ): Developing Mathematical Competence through Online-Assessment. *International Journal of Scholarship for Teaching and Learning (IJSTL), 15 (1,8).*

[6] Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. In R. A. Dienstbier (Ed.), *Nebraska Symposium on Motivation, 1990: Perspectives on motivation* (pp. 237-288). University of Nebraska Press.

[7] Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105* (1), 98-121.

[8] Foster, C. (2016). Confidence and competence with mathematical procedures. *Educational Studies in Mathematics*, 91 (2), 271-288.

[9] Foster, C. (2021). Implementing confidence assessment in low-stakes, formative mathematics assessments. *International Journal of Science Mathematics Education.*

[10] Gardner-Medwin, T. (2019). Certainty-based marking: Stimulating thinking and improving objective tests. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education: A handbook for academic practitioners* (2nd ed., pp. 141-150). Routledge.

[11] Gardner-Medwin, A. R. (2006). Confidence-based marking: Towards deeper learning and better exams. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp. 141-159). London: Routledge.

[12] Hassmén, P., Hunt, D. P. (1994). Human self-assessment in multiple-choice testing. *Journal of Educational Measurement*, 31, 149–160.

[13] Hudson, J., Bloxham, S., den Outer, B., & Price, M. (2017). Conceptual acrobatics: talking about assessment standards in the transparency era. *Studies in Higher Education*, 42 (7), 1309-1323.

[14] Lindquist, E. F., & Hoover, H. D. (2015). Some notes on corrections for guessing and related problems. *Educational Measurement: Issues and Practice*, 34 (2), 15-19.

[15] Maxwell, G.S. (2021). *Using Data to Improve Student Learning: Theory, Research and Practice.* Springer.

[16] Milner-Bolotin, M. (2018). Evidence-based research in STEM teacher education: From theory to practice. *Frontiers in Education* (3).

[17] Momsen, J.L., Long, T.M., Wyse, S., & Ebert-May, D. (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE Life Sciences Education* 9 (4), 435-440.

[18] Nathaniel, E.D, Scott, H.F, Wathen, B., Schmidt, S.K., Rolison, E., Smith, C., Hays, M.J., & Lockwood, J.M. (2021). Confidence-weighted Testing as an impactful education intervention within a pediatric sepsis. Quality Improvement Initiative. *Pediatric Quality and Safety*, 460.

[19] Newmark, B. (2019). *Why Teach*? John Catt Educational Ltd.

[20] Novacek, P. F. (2017). Exploration of a Confidence-Based Assessment Tool within an Aviation Training Program. *Journal of Aviation/Aerospace Education & Research, 26* (1).

[21] Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment and Evaluation in Higher Education*, 35 (5), 535-550.

[22] Sparck, E., Bjork, E., & Bjork, R. (2016). On the learning benefits of confidence-weighted testing. *Cognitive research: Principles and implications*, 1 (3).

[23] Stiggins, R. J. (2002). Assessment crisis: The absence of Assessment for Learning. Phi Delta Kappan, 83 (10), 758-765.

[24] Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46, 186-204.

[25] Wiliam, D. (2011). What is Assessment for Learning? *Studies in Educational Evaluation*, 37 (1), 3-14.

[26] Wu, Q., Vanerum, M., Agten, A., Christiansen, A., Vandenabeele, F., Rigo, JM, & Janssen, R. (2021). Certainty-Based Marking on Multiple-Choice Items: Psychometrics Meets Decision Theory. Psychometrika, 86 (2), 518-543.

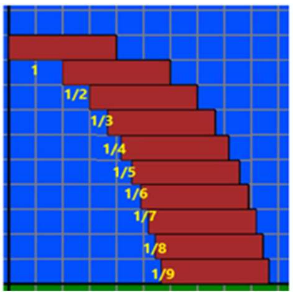
[27] Yen, Y. C., Ho, R. G., Chen, L. J., Chou, K. Y., & Chen, Y. L.(2010). Development and evaluation of a confidence-weighting computerized adaptive testing. *Educational Technology & Society*, 13 (3),163-176.

Appendix A
Student Generated Questions (SGQ)

1. The sequence of the partial sums of a series $\sum_{i=1, \dots, n} a_i$ is defined as the x-coordinate of the lower left corner of the n-th brick of a tower (brick 0 is at the top and lies at x = 0). For an element of the sequence $(a_n)_{n \in \mathbb{N}}$, a_i is the difference between the coordinates of the i-th and the i–1-th brick. Which of the following statements is / are correct?

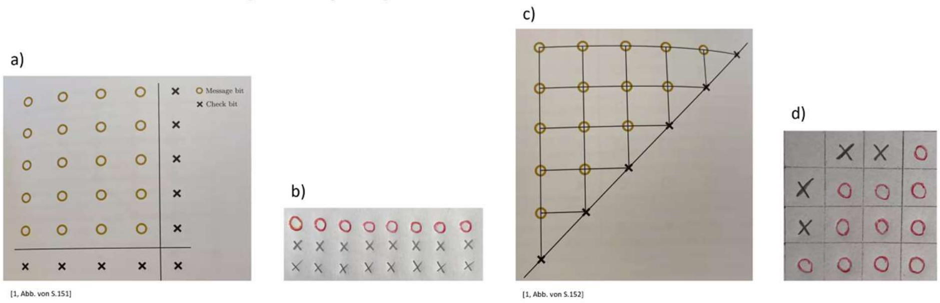
[Assumption: The size of the brick remains unchanged.]

a) If you can build a tower with an infinitely large ledge based on the partial sums, the series diverges.
b) If the tower topples over, the series diverges.
c) If you can build a tower, the series converges.
d) If the series converges, you can build a tower.



2. Which of the following approaches is the most robust one with regard to error correction? - Order them from the most to the least robust.

a) Order c,d,a,b
b) Order d,c,a,b
c) Order d,a,c,b
d) Order d,b,a,c



[1, Abb. von S.151]

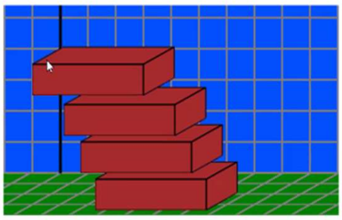
[1, Abb. von S.152]

3. Magic Mike says: “After the member of the audience has shuffled the cards if it can absolutely happen that 2 hearts or also 3 black cards are lying together. But this is no problem!”

Which of the following answers to Magic Mike’s comments would be correct?
a) “As the cards are always presented block by block, the audience wouldn’t notice.”
b) “The fact that k cards leave the remainders {0,…,k-1} when dividing by k means that there are k different cards in each of the blocks shown.
c) “You didn't get the trick because …”
d) “It cannot be that 3 black cards are next to each other– but it can certainly happen with features that appear in more than two variations (for example card value).”

4. Is it possible to build an infinite ledge in Two directions on the tower?

a) Yes, because the coordinates of the barycenter can be calculated separately for every coordinate direction.
b) No, the tower topples over.
c) Yes, if the corner point lies exactly under the barycenter of the tower on top.
d) No, because the downwards shift also influences the horizontal coordinate of the barycenter.



5. Is it possible to distort an image at 360° for a cylindrical mirror?
a) Yes, the image will be brought to focus at the front of the mirror anyway.
b) No, only works for 2 images.
c) No, there is no way to get the image back in focus.
d) Will not work with AnamorphMe, but can be done using grids.

6. Can we have more than two image distortions on the same anamorphic plane? For example, is it possible to distort 3 or 4 images, to be projected on the same cylindrical mirror?
   a) Yes, then the images will be close together on the mirror.

b) Yes, though the images will overlap on the mirror.
c) Depends on method of distortion being used.
d) No, we can only have a maximum of 2 images.

7. Which of the following properties is true for the extended Hamming Code?
a) It detects all errors, but it can only correct one of them.
b) It detects all even errors and can correct one bit if the error is a single error.
c) It detects all errors and can correct all even errors.
d) If there are an odd number (larger than 1) of errors, then neither the error detection nor the error correction works.

Appendix B

Task Perception Questionnaire (TPQ)

1. Certainty Based Marking was a relevant and useful activity.
(a) Strongly agree (b) agree (c) disagree (d) strongly disagree

2. Certainty Based Marking was mentally very demanding.
(a) Strongly agree (b) agree (c) disagree (d) strongly disagree

3. Certainty Based Marking made me think deeper (more reflective).
(b) Strongly agree (b) agree (c) disagree (d) strongly disagree

4. I will continue producing questions when I learn new materials in the future.
(a) Strongly agree (b) agree (c) disagree (d) strongly disagree

5. I will continue re-assessing my answers to become more confident.
(b) Strongly agree (b) agree (c) disagree (d) strongly disagree