Article

Student Reactions to Just-in-Time Formative and Summative Feedback in a Tablet-Based Family Medicine MCQ Exam

Johanna Klutmann, Constanze Dietzsch, Philip Vogt, Nadine Wolf [ID], Melanie Caspar, Sara Volz-Willems [ID], Johannes Jäger and Fabian Dupont *[ID]

Department of Family Medicine, Saarland University, 66421 Homburg, Germany
* Correspondence: fabian.dupont@uks.eu

Abstract: (1) Background: While the benefits of digital assessments for universities and educators are well documented, students' perspectives remain underexplored. (2) Methods: This study employed an exploratory mixed-methods approach. Three examination cohorts were included (winter semester 2022/23, summer semester 2023, and winter semester 2023/24). Written emotional responses to receiving just-in-time summative or formative feedback were analyzed, as well as examining the impact of formative feedback on learning attitudes. All cohorts completed qualitative open-ended research questions. The responses were coded using Kuckartz's qualitative content analysis. Descriptive statistics were generated using jamovi. (3) Results: Students generally responded positively to formative and summative feedback. The majority expressed a desire to receive feedback. The categories created for formative feedback indicate a tendency toward self-reflection and supported the learning processes. In contrast, the summative feedback categories suggest that students primarily value feedback's transactional aspect. (4) Conclusions: Integrating formative and summative feedback in digital just-in-time assessment offers the potential to capitalize on the "sensitive periods" of study reflection that occur during assessment. This approach enables assessment for learning while simultaneously reducing emotional distress for students.

Keywords: medical education; formative feedback; summative feedback; digital assessment

1. Introduction

A significant amount of research in medical education (ME) has been dedicated to the examination of assessment practices and the potential for their enhancement [1]. Multiple-choice questions (MCQs) are a prevalent assessment tool in ME [2]. Critics have highlighted the reliance on well-trained faculty to develop MCQs that assess higher-order thinking (HOT) [3,4].

In Germany, MCQs constitute an integral component of the ME curriculum, as the written state examinations are exclusively based on MCQs [5]. With the aim of supporting ongoing MCQ development, this study has drawn upon key research findings in the field of medical assessment, focusing on four areas: digitalization, key-feature questions (KFQs), formative feedback (FF), and summative feedback (SF).

The utilization of digitally supported assessments is increasing in ME. The digitization of assessments offers numerous advantages for educators [6]. The integration of media such as images, audio, and video into clinical cases or KFQs is facilitated [7]. KFQs are short clinical cases followed by questions on essential decision-making elements. They make

examinations more contextual and allow for the assessment of clinical decision-making skills without changing the basic format of an MCQ assessment [8].

In a digital format, KFQs can be locked, thereby preventing students from returning to previous questions after receiving additional information or feedback. This digital feature enables the precise analysis of students' responses. It demands that students engage in the application and evaluation of concepts to answer the questions. The design of this digital feature therefore promotes HOT. The concept of HOT is derived from the Bloom taxonomy of the cognitive domain in 1956 [9]. The cognitive domain comprises hierarchical levels of thinking. These levels have been updated to include remembering, understanding, applying, analyzing, evaluating, and creating [10].

FF is a core element of "assessment *for* learning" [11]. It provides feedback during the learning process with a focus on assisting learning [11]. FF is known to further enhance learning. Students value and desire it because it influences their study habits and deepens their understanding, even during assessments [12]. This study explores the broader implication of FF, extending it beyond the scope of a solitary assessment. Final examinations are not regarded as the culmination of a learning activity, but rather as a subsequent opportunity for further learning. The provision of feedback assumes a formative character, enabling students to recalibrate their learning strategies over the course of time.

SF, which concentrates primarily on the "assessment *of* learning", evaluates what has been learned [11]. Although it does not always promote a deeper understanding of the assessed topics, it is considered a driving force of learning due to its importance [13].

Certain qualities of feedback enhance students' receptivity to it. These include: autonomy in assessment, the student's perception of its relevance, the clarity of a grade as a reassuring aspect, and constructive advice to complement negative feedback [14,15].

This study emerged in response to a shift in German education policy towards competency-based medical education (CBME). This included the aim for competency-based and digital teaching, strengthening Family Medicine and improving assessments [16,17]. When considering how to improve feedback in assessments, the timing of feedback emerges as a critical factor. In particular, post-assessment and immediate feedback is viewed as highly important and effective [18,19]. Not only does it trigger cognitive processes at a receptive time, thereby supporting learning, but turns a pure assessment into a 'test [that] would teach as well as assess' [20]. During the tablet-based KFQ assessments, just-in-time feedback and the research questions could be provided conveniently and immediately. This enabled the accurate capture of students' motivations and emotions. This process constitutes an evaluation of teaching effectiveness, a key component of CBME [21], which has not been executed in this format previously.

Following Epstein's assessment goals, this study focused on the first goal: uncovering the motivations behind students' desire for SF and the feelings evoked by its potential just-in-time delivery [22]. It also compared these emotions to those elicited by FF.

In collaboration with the German Institute for State Examinations in Medicine, Pharmacy, Dentistry, and Psychotherapy (IMPP, Institut für medizinische und pharmazeutische Prüfungsfragen) and Saarland University, just-in-time FF and SF were incorporated into the tablet-based MCQ assessment at the end of the Family Medicine curriculum in the fifth year of medical school. The integration process was executed with two overarching objectives in mind. Firstly, to provide a comprehensive description of the capabilities of MCQs. Secondly, it sought to acquire novel insights regarding the potential of MCQs to facilitate learning. This study examined the impact of the implementation of tablet-based, just-in-time feedback on the motivation, emotional engagement, and learning attitudes of medical students within a competency-based medical education framework in Germany.

2. Materials and Methods

2.1. Setting

An exploratory mixed-methods study was conducted, incorporating a qualitative and quantitative component in the data analysis. The qualitative component of the study included research questions exploring students' reactions to FF and SF and the emotions evoked by them. The quantitative component of the study employed a descriptive analysis to ascertain the distribution of categories. The study encompassed three data collections. The initial data collection occurred during the winter semester (WS) of 2022/2023 (WS 22/23). The second data collection occurred during the summer semester (SS) of 2023 (SS 23), and the third and final data collection took place in the winter semester (WS) of 2023/24 (WS 23/24). All data collections took place during the final tablet-based exam of the Family Medicine curriculum at the end of the fifth year of medical school at Saarland University. All students were already familiar with the tablet due to previous tablet-based exams.

2.2. Study Participants

In WS 22/23, 97 out of 99 students participated. In SS 23, 115 out of 115 students participated. In WS 23/24, 95 out of 99 students participated. All participants were undergraduate medical students officially enrolled in the Family Medicine curriculum at Saarland University. The exclusion criteria were as follows: not consenting to participate or skipping responses to the research questions. Consent to participate was sought during the two weeks before the exam. Ethical approval was obtained by the responsible ethics committee for medical studies at Saarland University (234/20-14 April 2022).

2.3. Data Collection

The exams consisted of 60 constructively aligned MCQs and included an additional research component after the main exam. In the first study arm (WS 22/23), the research section allowed students to see preliminary SF in the form of scores immediately after completing the exam. This was the first time this option had been offered to these students. As can be seen in Figure 1, they were asked about their reasons and their feelings regarding this option.
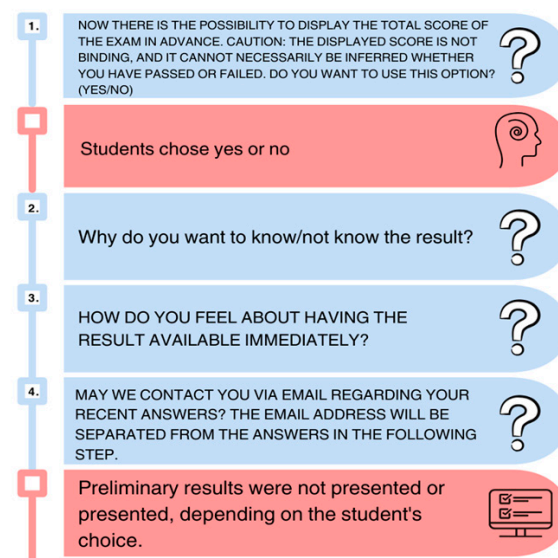

Figure 1. First study arm, including research questions and the procedure of the WS 22/23 exam, with the provision of just-in-time summative feedback at the end.

In the second study arm (SS 23), the additional research component included two KFQs. The first KFQ comprised two sub-questions (1.1 and 1.2). The second KFQ comprised four sub-questions (2.1–2.4). The individual questions were comparable to those of the main examination, as they were adapted from official IMPP questions. Participants were randomly divided into two groups: Group A (*n* = 63) and Group B (*n* = 52). Group A received FF in the form of the correct answer with an added explanation after each of the four sub-questions of the second KFQ (after questions 2.1–2.4). Conversely, Group B received FF subsequent to each of the two sub-questions of the first KFQ (after questions 1.1 and 1.2). Following the provision of FF, students were prompted to share their feelings and perceptions regarding the received feedback (Figure 2).
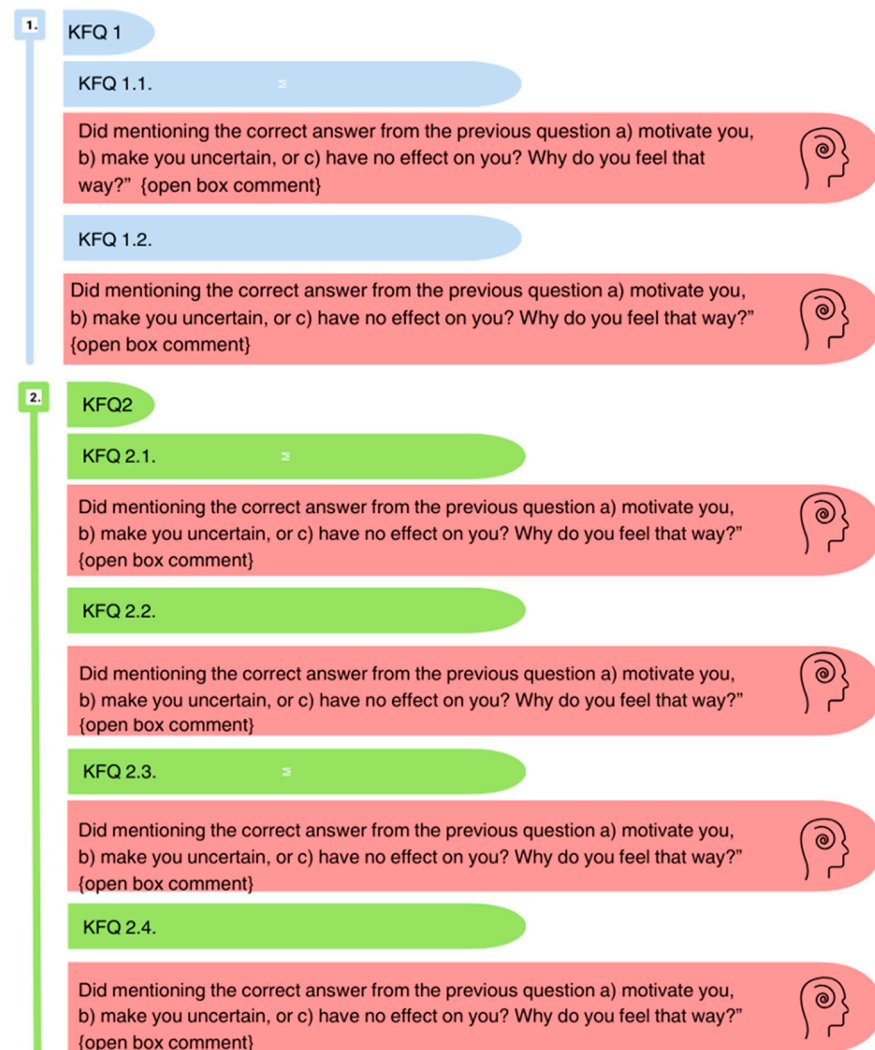

Figure 2. Second study arm, including research questions and the procedure of the SS 23 exam, with the provision of just-in-time formative feedback after each question.

In the third study arm (WS 23/24), the additional research component included two KFQs with sub-questions (1.1, 1.2, 2.1, 2.2, 2.3, and 2.4). Participants were divided into two groups: Group A (*n* = 44) and Group B (*n* = 54). Following the completion of one of the KFQs, the participants received FF. At the conclusion of the research component, the students were asked about their preferred timing for feedback and whether they believed the FF would influence their learning attitude (Figure 3).

Figure 3. Third study arm, including research questions and procedure of the WS 23/24 exam, including questions about the preferred timing of feedback and the influence of FF on learning attitude.

2.4. Data Analysis

Kuckartz's model of qualitative content analysis was employed in the analysis of the students' written comments [23]. To ensure qualitative credibility, a researcher triangulation session was organized after each data collection (FD, JJ, SVW, MC, JK, and CD). First, separate coders (JK and CD for WS 22/23, JK, CD, and PV for SS 23, and JK, PV, and NW for WS 23/24) categorized the data into main categories. The same categories were identified by four additional researchers (FD, JJ, SVW, and MC). Second, the coders engaged in a collaborative discussion to create inductive subcategories. The students' responses occasionally encompassed multiple statements. In such instances, they were categorized into multiple categories (non-exclusivity). The qualitative data were then quantitatively analyzed with descriptive statistics as a percentage within the cohort, using jamovi version 2.4 12.0.

3. Results

A total of 310 students participated (WS 22/23: 97, SS 23: 115, and WS 23/24: 95). Exam performance was comparable for the three semesters, with average student scores of 80–85%.

3.1. Just-in-Time Summative Feedback

3.1.1. Students' Motivation for Just-in-Time Summative Feedback

In the first study arm, which involved the provision of just-in-time SF, 95 out of 97 students elected to receive SF. The most common motivation was to actively *remove negative emotions* (37.5%) (e.g., 'It makes me worry less' {SF.10.1}). Other motivations included *self-confirmation* (28.1%). This encompassed the confirmation of passing the exam or the student's familiarity with the knowledge being assessed (e.g., '[It gives me] direct comparison with the self-assessment after the exam and [I] know directly whether [my feeling] was right' {SF.49.1}). Students wanted to plan their exam and study time (16.7%), which aligned with the theme of *planning security* (e.g., '[I would like to know the result, because of the possible] re-examination and whether [I should] study again or not concerning schedules with other exams' {SF.22.1}). Some students (14.6%) were simply curious about the results i.e., out of *interest* (e.g., '[I do not] fear the result, I just want to know it' {SF.41.1}). Another motivation was *peer pressure* (2.1%). Students felt compelled to see the SF because they knew it would be a topic of conversation after the exam (e.g., 'Because otherwise, everyone else will know before I do, but I prefer to have a bit of time before the results are announced' {SF. 20.1}). Students wanted to find closure in the exam without providing an additional emotional context (12.5%). This was coded as *closure, not further emotionally described* (e.g., 'So [I] know as quickly as possible to what extent [my] inner feeling is correct or not' {SF.32.1}).

3.1.2. Motivation Against Just-in-Time Summative Feedback

A small percentage of students (2.1%) did not want to see the SF, due to *fear of being demotivated* during the upcoming exams (e.g., 'Fear of a bad result. It would inhibit my studying for the upcoming exams' {SF.33.1}).

3.1.3. Emotional Responses Following Just-in-Time Summative Feedback

Students were asked how they felt about receiving just-in-time SF, just before receiving it. Most students (67.7%) reported a *positive emotion*. Relief was a common theme (e.g., 'Pleasant, helpful' {SF. 21.2}, 'Relief, no matter how it turns out'{SF.44.2}). Some students (5.2%) reported a *negative emotion* (e.g., 'Unpleasant' {SF.33.2}). Others (10.4%) reported feeling *nervous* (e.g., 'It makes me nervous' {SF.61.2}). For 10.3% of the students, the emotion was *ambivalent* (e.g., 'Exciting and reassuring at the same time. Paradoxical, isn't it?' {SF.13.2}).

Some participants (12.5%) did not answer the research question.

Two themes were identified regarding the perception of just-in-time SF. For 7.3%, SF was *outcome-dependent* (e.g., 'On one hand good if you had a good feeling and on the other hand it can ruin your day if you didn't do well' {SF.49.2})**.** Second, there was the importance of being able to voluntarily opt for SF (5.2%). A desire for *autonomy regarding SF* was expressed (e.g., 'I think you should have the possibility to see the score, but you shouldn't force anyone'{SF.6.2}).

3.2. Just-in-Time Formative Feedback

3.2.1. Motivation for Just-in-Time Formative Feedback

The third study arm (WS 23/24) asked for students' preferred timing of feedback. Most students (83.1%) wanted to receive just-in-time FF. Students (38.9%) saw FF as a *stress reduction*. (e.g., '[FF] provides relief from other uncertainties and clears the mind for the next question' {LB.93.1}). Students experienced a *learning effect* (36.8%). By better understanding the topic or avoiding mistakes, students gained new knowledge from FF (e.g., '[I want to receive] feedback as quickly as possible, [as it] enhances [the] learning process' {LB.34.1}). Students perceived just-in-time FF as *self-confirmation* (8.4%). It strengthened their self-confidence and allowed them to trust their previous approach to answering questions (e.g., 'This allows me to quickly confirm my thought process. I can better assess whether I am on the right track or not and can avoid making mistakes later' {LB.35.1}). A few students (3.2%) were *curious* about just-in-time FF (e.g., 'Because I would just like to know instantly whether I have answered it correctly' {LB.43.1}, 'curiosity' {LB.32.1}). The least number of students (1.1%) reported *motivation for the exam* as their reason (e.g., '[FF] motivates me to answer the next question' {LB.40.1}).

3.2.2. Motivation Against Just-in-Time Formative Feedback

Overall, 16.9% of the students were opposed to FF in an MCQ exam. Some students (7.4%) preferred to receive feedback during the exam review, a few weeks after the exam. Others (9.5%) desired a different time, most commonly the evening of the exam. Providing just-in-time FF generated negative feelings in some students. A portion of the students *feared uncertainty* from the thought of possibly having answered incorrectly and finding out about it during the exam (8.4%) (e.g., '[A] wrong answers in the middle of the exam would only unsettle me and throw me off' {LB.33.1}). Students found it harder to concentrate during the exam, viewing just-in-time FF as a *distraction* (5.3%) and desiring an exam without interruptions (e.g., '[I would like to have] no distractions [during the exam], and I still remember the question [during the exam review]' {LB.6.1}). Some students *desired relaxation* (3.2%). The prospect of receiving FF was perceived to be a source of stress (e.g., 'So that you don't put yourself under so much pressure during the exam' {LB.22.1}).

3.2.3. Emotional Responses Following Just-in-Time Formative Feedback

In the second study arm (SS 23), the provision of just-in-time FF generated three main categories: *motivation*, *uncertainty*, and *no effect*. Some responses did not fit into the main categories, necessitating the addition of the category *frustration*.

Most students felt motivated (74.2%), 18% felt no effect, and 7% felt uncertain after receiving FF. A minority of 1.1% reported frustration.

Motivation as a category was further subcategorized. When students answered correctly, 12.9% felt a *sense of accomplishment* (e.g., 'Because I answered the question correctly' {FF. 1.71}). Others reported *reassurance* (5.1%). The uncertainty felt when answering the question was alleviated by the feedback (e.g., '[motivated], because I was unsure here' {FF.5.21}). Regardless of whether the students gave the right or wrong answer, FF was perceived by 2.2% as an *opportunity for a learning process* (e.g., '[Because of FF] I got a detailed explanation and new information' {FF.6.17}). Students reported a feeling of *self-confirmation* after receiving FF (20.2%). This feeling can be divided into four subcategories.

1. Confirmation of learning method (2.5% of all students)

FF was seen as a confirmation of individual learning methods. The time and effort the students invested was seen as worthwhile (e.g., '[It is a] feeling of confirmation [and shows that] I have studied well' {FF.4.30});

2. Confirmation of thought process (6.2% of all students)

The FF confirmed students' thought processes or clinical reasoning. They felt validated in continuing to use it in the future (e.g., '[motivated] as my thought process regarding my choice of therapy was confirmed' {FF.1.73});

3. Confirmation of intuition (2.8% of all students)

The FF confirmed the students' intuition. Sometimes, students relied on a gut feeling rather than a logical thought process (e.g., '[motivated], because my gut feeling was right' {FF.5.45)};

4. Confirmation of Existing Knowledge (3.9% of all students)

The students confirmed their prior knowledge with the FF (e.g., 'Since I am sure, it motivated me to stand up for my decision and justify it' {FF.3.37}).

Uncertainty

The reasons for students' uncertainty after receiving FF were a *lack of understanding of the task* and *incorrect knowledge*. Some participants (0.3%) reported uncertainty due to wording (e.g., 'I don't know what you mean by 'motivated' {FF.1.109}). Due to incorrect knowledge or a perceived gap in their own knowledge regarding the assessed topic, 3.1% of students felt uncertain (e.g., '[I feel uncertain], because of a knowledge gap' {FF.5.34}).

No Effect

Some students reported that FF did not affect them (16.9%). A few students (7.3%) were *confident in their knowledge* and response prior to FF (e.g., '[I am] not uncertain because I was very sure' {FF.3.49}). A minority of 0.8% of students knew that their answer was wrong before receiving FF (e.g., 'I was already unsure about the answer anyway' {FF.4.40}).

Frustration

Answering the question incorrectly caused 1.1% of all students (4,5% of those who answered incorrectly) to feel frustrated (e.g., 'It frustrates me because I quickly realized that I made the wrong choice' {FF.4.7}). On the other hand, frustration also occurred in the context of motivation (e.g., '[motivation], but it frustrates me because I should have known' FF.4.38}).

3.2.4. Effect of Just-in-Time Formative Feedback on Learning Attitudes

In the third study arm (WS 23/24), when asked if the just-in-time feedback would influence their learning attitudes, most students agreed (75.8%), while a few disagreed (10.5%). Most students reported a positive influence on their learning attitude (66.3%), while others reported a negative one (9.5%).

Positive Influences

Most students (23.2%) used the just-in-time FF as an *error correction* (e.g., 'I deal with it directly, with a corresponding explanation. If my answer is wrong, now I have the correct answer with an explanation' {LB.51.2}). *Memorization* (22.1%) was also an important theme. Due to the just-in-time FF, students felt more likely to remember the feedback and recall it (e.g., 'This helps me to study, I memorize the learning content better' {LB.81.2}). Some students reported *linking* (12. 6%). The just-in-time FF helped students link new content with previously known knowledge (e.g., 'I can better recognize connections and think about them' {LB.94.2}). Just-in-time FF evoked *concentration* (6.3%). Students gained a better focus for the exam (e.g., 'It sharpens my concentration a little more . . . so that I read more attentively for the next question' {LB77.2}). Correct responses were perceived as *motivating*

(4.2%). This perception is associated with both motivation for the examination and learning in general (e.g., 'Becoming more motivated to learn' {LB.9.2}). Some students reported a *learning reflection* (4.2%). The new information encouraged students to examine their previous knowledge or learning strategies (e.g., 'I can question my clinical thinking directly' {LB.89.2}). FF addressed gaps in the students' knowledge (3.2%). Students were informed about which topics they needed to revise, immediately *filling knowledge gaps* (e.g., 'Direct feedback sticks better, personal gaps are better visualized' {LB.5.2}).

Negative Influence

The just-in-time FF was seen as *unsettling* (6.3%), as it highlighted the number of incorrect answers. Some students noted a personal difference between an exam setting and a study-only setting, in which they would prefer just-in-time FF (e.g., 'I think it is unsettling during the exam' {LB.19.2}). Students felt *distracted* by the just-in-time FF (2.1%), leading them to ruminate on their mistakes and lose concentration (e.g., 'Yes, [a] distraction and [I'm] thinking about mistakes made. [I have] less concentration' {LB.6.2}).

4. Discussion

The results indicate that most students on the course of Family Medicine appreciate just-in-time FF and SF. This study yielded three main findings: First, most students expressed a desire for just-in-time feedback and a motivation to engage with it. Second, the categories created for FF indicate self-reflection and learning processes. The SF categories show that students still want a straightforward conclusion without always seeking a deeper understanding of their mistakes. Third, most students agreed with the positive influence of just-in-time FF on their learning attitude.

The reception of SF was positive for 67.7% of the participating students. The strongest motivational responses to SF were *removing negative emotions* and *self-confirmation*. These categories can be interpreted as behavioristic, as the desire to eliminate negative stimuli and enhance positive stimuli seem to be key motivating effects triggered by SF. The SF categories also demonstrated the students' desire for a straightforward conclusion. *Removing negative emotions*, *planning security*, and *closure* exemplify students who conclude with the assessments without seeking an understanding of their errors.

Most students reported feeling motivated after receiving FF. There was a clear difference in the diversity and richness of qualitative responses to FF compared to SF. However, only 2.2% of students explicitly stated that they perceived FF as a learning experience in SS 23. In contrast, the FF categories suggest that a considerable proportion of students engaged in introspection regarding their learning. The students sought validation for various aspects of their learning (i.e., *intuition*, *thought process*, *existing knowledge*, and *learning method*). Moreover, the results in WS 23/24 demonstrated that the majority of students agreed that the just-in-time FF would exert a positive influence on their learning attitudes. Consequently, the findings suggest that the responses to FF indicated an aspiration for deeper reflection on learning and comprehension of the subjects addressed in the assessment. This aligns with previous findings and demonstrates that integrating FF in an assessment setting is broadly accepted by students and can be seen as a separate learning activity [24]. Although this is only a subjective opinion of their learning attitudes, the students' responses suggest the clear effect of FF on them. It shows their tendency to reflect on their learning in an MCQ assessment. Follow-up research should further investigate how this type of FF affects the learning attitudes of students. The findings reveal contrasting results in terms of engagement with feedback between FF and SF and its influence on learning attitudes. This discrepancy may be due to the nature of the two feedback methods. An assessment

accompanied solely by SF might not adequately motivate further engagement. In contrast, FF provides a structured framework and motivation for students to reflect on their errors.

With digital assessment, the risk of influencing performance (making the exam easier) can be avoided by locking questions (not allowing revisits). For educators, it may mean the possibility of combining learning activity and reflection with assessment. This could entail two distinct approaches: First, the utilization of FF in alternative learning activities, such as practice tests. This approach is substantiated by the findings of the present study, which demonstrated that students positively valued SF and FF. Secondly, viewing one assessment as a puzzle piece. This approach could offer students just-in-time feedback on their errors, thought processes, and recommendations for improvement over time. Final examinations would become a subsequent opportunity for further learning. Testing, as opposed to studying, can increase students' receptivity to information and improve retention [25,26]. This makes assessments a sensitive phase and students more likely to integrate feedback into their learning behavior. This is also evident in the categories that describe the positive impacts of just-in-time FF on learning attitudes (*learning reflection*, *memorization*, and *linking*). Combined with the sense of confirmation and the provision of planning security for the ongoing study period, these factors seem to support the inclusion of both FF and SF voluntarily in future digital MCQ assessments.

The data were collected just-in-time during the exam, effectively capturing the atmosphere of a real exam situation. The FF setup imitates the feedback structure of a popular and widely used digital learning platform in Germany—AMBOSS [27]. The overall participation rate was 99% across all semesters, reducing the likelihood of selection bias. This study offers novel insights into the use of just-in-time FF, as it can trigger learning reflection. This has been difficult to implement in traditional academia, especially in MCQ exams. The digital just-in-time SF and FF may present an opportunity to integrate structured feedback into future assessments in grade-oriented countries. This study provides evidence for the use of FF, a highly researched teaching tool, in an assessment environment. It thereby encourages and fosters the concept of "assessment for or as learning".

Limitations

The study design encompasses three sequential research questions. Generalized conclusions from one cohort to others are therefore not possible. Furthermore, this was the course of Family Medicine, i.e., only one course in the entire medical training program. In addition, the fifth year of study is a special situation, as preparations for the second state examination follow. This may enhance possible openness for rapid clarification. The overall exam performances were high (average scores of 80–85%). This may account for the relatively low percentage of students who expressed uncertainty about receiving just-in-time FF. Future studies should aim to explore how categories may differ with a broader variation in difficulty levels, particularly during the earlier stages and different departments of undergraduate ME. FF in an MCQ assessment is not comparable to verbal FF in a one-on-one conversation in terms of individuality, but this study showed the subjective positive impact this format can have on students.

5. Conclusions

Students' positive reactions to FF and SF indicate an open-mindedness to change and a willingness for the further development of exam structures regarding feedback implementation. This creates additional opportunities for educators. Students desire more than SF. The combination of FF and SF enables the leveraging of sensitive periods of reflection on learning during an exam that may be difficult to facilitate otherwise. Ideally, reflection may alter future learning attitudes. Using just-in-time feedback in digital

assessments is time-efficient and provides a valuable way to integrate feedback in grade-oriented countries like Germany. In this way, an exam would remain a reliable summative assessment (as is often a prerequisite), while also becoming a significant learning and reflection activity for students (assessment as learning).

Abbreviations

The following abbreviations are used in this manuscript:

ME	Medical education
MCQ	Multiple-choice questions
KFQ	Key-feature question
HOT	Higher-order thinking
IMPP	Institut für medizinische und pharmazeutische Prüfungsfragen (Institute for State Examinations in Medicine, Pharmacy, Dentistry, and Psychotherapy (IMPP))
WS	Winter semester
SS	Summer semester
FF	Formative feedback
SF	Summative feedback
CBME	Competency-based medical education

References

1. Holmboe, E.S.; Sherbino, J.; Long, D.M.; Swing, S.R.; Frank, J.R. The role of assessment in competency-based medical education. *Med. Teach.* **2010**, *32*, 676–682. [CrossRef] [PubMed]
2. Grainger, R.; Dai, W.; Osborne, E.; Kenwright, D. Medical students create multiple-choice questions for learning in pathology education: A pilot study. *BMC Med. Educ.* **2018**, *18*, 201. [CrossRef]

3. Javaeed, A. Assessment of Higher Ordered Thinking in Medical Education: Multiple Choice Questions and Modified Essay Questions. *MedEdPublish* **2018**, *7*, 128. [CrossRef]

4. Vanderbilt, A.A.; Feldman, M.; Wood, I. Assessment in undergraduate medical education: A review of course exams. *Med. Educ. Online* **2013**, *18*, 20438. [CrossRef] [PubMed]

5. Chenot, J.-F. Undergraduate medical education in Germany. *GMS Ger. Med. Sci.* **2009**, *7*, Doc02. [CrossRef]

6. Thoma, B.; Turnquist, A.; Zaver, F.; Hall, A.K.; Chan, T.M. Communication, learning and assessment: Exploring the dimensions of the digital learning environment. *Med. Teach.* **2019**, *41*, 385–390. [CrossRef]

7. Egarter, S.; Mutschler, A.; Tekian, A.; Norcini, J.; Brass, K. Medical assessment in the age of digitalisation. *BMC Med. Educ.* **2020**, *20*, 101. [CrossRef] [PubMed]

8. Al-Wardy, N.M. Assessment methods in undergraduate medical education. *Sultan Qaboos Univ. Med. J.* **2010**, *10*, 203. [CrossRef]

9. Forehand, M. Bloom's taxonomy. *Emerg. Perspect. Learn. Teach. Technol.* **2010**, *41*, 47–56.

10. Krathwohl, D.R. A revision of Bloom's taxonomy: An overview. *Theory Into Pract.* **2002**, *41*, 212–218. [CrossRef]

11. Murugesan, M.; David, P.L.; Chitra, C.B. Correlation between Formative and Summative Assessment Results by Post Validation in Medical Undergraduates. *IOSR J. Dent. Med. Sci. (IOSR-JDMS) E-ISSN* **2021**, *20*, 51–57. [CrossRef]

12. Badyal, D.K.; Bala, S.; Singh, T.; Gulrez, G. Impact of immediate feedback on the learning of medical students in pharmacology. *J. Adv. Med. Educ. Prof.* **2019**, *7*, 1. [CrossRef] [PubMed]

13. O'Shaughnessy, S.M.; Joyce, P. Summative and Formative Assessment in Medicine: The Experience of an Anaesthesia Trainee. *Int. J. High. Educ.* **2015**, *4*, 198–206. [CrossRef]

14. Harrison, C.J.; Könings, K.D.; Dannefer, E.F.; Schuwirth, L.W.; Wass, V.; van der Vleuten, C.P. Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. *Perspect. Med. Educ.* **2016**, *5*, 276–284. [CrossRef]

15. Hill, J.L.; Berlin, K.; Choate, J.K.; Cravens-Brown, L.; McKendrick-Calder, L.A.; Smith, S.F. Exploring the Emotional Responses of Undergraduate Students to Assessment Feedback: Implications for Instructors. *Teach. Learn. Inq.* **2021**, *9*, 294–316. [CrossRef]

16. Medizinischer Fakultätentag. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM). 2015. Available online: https://medizinische-fakultaeten.de/wp-content/uploads/2021/06/nklm_final_2015-12-04.pdf (accessed on 15 January 2025).

17. Medizinischer Fakultätentag. Medizinische Fakultäten Veröffentlichen die Neufassung des Nationalen Kompetenzbasierten Lernzielkatalogs. 2021. Available online: https://medizinische-fakultaeten.de/medien/presse/medizinische-fakultaeten-veroeffentlichen-die-neufassung-des-nationalen-kompetenzbasierten-lernzielkatalogs/ (accessed on 15 January 2025).

18. Choi, S.; Oh, S.; Lee, D.H.; Yoon, H.S. Effects of reflection and immediate feedback to improve clinical reasoning of medical students in the assessment of dermatologic conditions: A randomised controlled trial. *BMC Med. Educ.* **2020**, *20*, 146. [CrossRef]

19. Lee, G.B.; Chiu, A.M. Assessment and feedback methods in competency-based medical education. *Ann. Allergy Asthma Immunol.* **2022**, *128*, 256–262. [CrossRef]

20. Epstein, M.L.; Lazarus, A.D.; Calvano, T.B.; Matthews, K.A.; Hendel, R.A.; Epstein, B.B.; Brosvic, G.M. Immediate Feedback Assessment Technique Promotes Learning and Corrects Inaccurate first Responses. *Psychol. Rec.* **2002**, *52*, 187–201. [CrossRef]

21. McGaghie, W. *Competency-Based Curriculum Development in Medical Education: An Introduction*; Public Health Papers; WHO: Geneva, Switzerland, 1978; No. 68.

22. Epstein, R.M. Assessment in medical education. *N. Engl. J. Med.* **2007**, *356*, 387–396. [CrossRef]

23. Kuckartz, U. *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung*; Beltz Juventa: Weinheim, Germany; Basel, Switzerland, 2012.

24. Kornegay, J.G.; Kraut, A.; Manthey, D.; Omron, R.; Caretta-Weyer, H.; Kuhn, G.; Martin, S.; Yarris, L.M. Feedback in medical education: A critical appraisal. *AEM Educ. Train.* **2017**, *1*, 98–109. [CrossRef]

25. Agarwal, P.K.; Karpicke, J.D.; Kang, S.H.; Roediger, H.L., III; McDermott, K.B. Examining the testing effect with open-and closed-book tests. *Appl. Cogn. Psychol. Off. J. Soc. Appl. Res. Mem. Cogn.* **2008**, *22*, 861–876. [CrossRef]

26. French, S.; Dickerson, A.; Mulder, R.A. A review of the benefits and drawbacks of high-stakes final examinations in higher education. *High. Educ.* **2024**, *88*, 893–918. [CrossRef]

27. Bientzle, M.; Hircin, E.; Kimmerle, J.; Knipfer, C.; Smeets, R.; Gaudin, R.; Holtz, P. Association of online learning behavior and learning outcomes for medical students: Large-scale usage data analysis. *JMIR Med. Educ.* **2019**, *5*, e13529. [CrossRef] [PubMed]