



Universiteit
Leiden
The Netherlands

Assessment for growth: fostering student learning through assessment innovations in medical education

Wijk, E.V. van

Citation

Wijk, E. V. van. (2025, November 19). *Assessment for growth: fostering student learning through assessment innovations in medical education*. Retrieved from <https://hdl.handle.net/1887/4283162>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4283162>

Note: To cite this publication please use the final published version (if applicable).



Part III: Feedback

Chapter 9

Understanding students' feedback use in medical progress testing: A qualitative interview study

Elise V. van Wijk
Floris M. van Blankenstein
Roemer J. Janse
Eline A. Dubois
Alexandra M.J. Langers

Medical Education. 2024;58(8):980-988.
DOI: 10.1111/medu.15378.

Abstract

Background Active engagement with feedback is crucial for feedback to be effective and improve students' learning and achievement. Medical students are provided feedback on their development in the progress test (PT), which has been implemented in various medical curricula, although its format, integration and feedback differ across institutions. Existing research on engagement with feedback in the context of PT is not sufficient to make a definitive judgement on what works and which barriers exist. Therefore, we conducted an interview study to explore students' feedback use in medical progress testing.

Methods All Dutch medical students participate in a national, curriculum-independent PT four times a year. This mandatory test, composed of multiple-choice questions, provides students with written feedback on their scores. Furthermore, an answer key is available to review their answers. Semi-structured interviews were conducted with 21 preclinical and clinical medical students who participated in the PT. Template analysis was performed on the qualitative data using a priori themes based on previous research on feedback use.

Results Template analysis revealed that students faced challenges in crucial internal psychological processes that impact feedback use, including 'awareness', 'cognizance', 'agency' and 'volition'. Factors such as stakes, available time, feedback timing and feedback presentation contributed to these difficulties, ultimately hindering feedback use. Notably, feedback engagement was higher during clinical rotations, and students were interested in the feedback when seeking insights into their performance level and career perspectives.

Conclusion Our study enhanced the understanding of students' feedback utilization in medical progress testing by identifying key processes and factors that impact feedback use. By recognising and addressing barriers in feedback use, we can improve both student and teacher feedback literacy, thereby transforming the PT into a more valuable learning tool.

Introduction

Effective feedback improves students' learning and achievement, fosters adaptive learning and prepares students for life-long learning [1-3]. The longitudinal medical progress test (PT) is implemented in various countries around the world [4]. Although PTs are operationalised in different ways across medical schools, their core function is to provide medical students feedback on their knowledge growth throughout their studies. This feedback is aimed at assessment *for* learning as it stimulates identification of learning gaps and adjustment of learning [1, 5-7]. Feedback is effective when students actively engage with it and act upon it. However, in practice, this rarely happens, which raises doubts about the effectiveness of PT feedback [8-13].

To promote student engagement with feedback, students need to acquire feedback literacy, that is, be able to understand, appreciate, utilise and benefit from feedback processes. This requires a proactive attitude and a shift towards a learning-centred approach [14-16]. Based on the framework of Carless and Boud, [14] Molloy *et al.* [15] identified seven essential categories of student feedback literacy, including the understanding of feedback purposes, emotional engagement, and utilising the feedback for future work. Furthermore, Winstone *et al.* [17] identified four pivotal internal psychological processes for effective feedback use: (1) *awareness* of what feedback means and what its purpose is, (2) *cognizance* of appropriate strategies to implement feedback, (3) *agency* in implementing these strategies and (4) *volition* (or will) to explore and act on the feedback. Difficulties in these processes form barriers to effective feedback use. In undergraduate medical education, feedback receptiveness is primarily influenced by students' characteristics (e.g. confidence and mindset), feedback content, educators' credibility and the learning environment [18]. Guidelines for effective feedback in clinical learning underscore the impact of clear and supportive feedback on motivating trainees. The most important elements include enhancing self-efficacy and fostering the development of strategies that lead to improved competencies [19]. Additionally, teachers and institutions play a crucial role in fostering a climate that supports learning and enhances engagement with feedback [16, 18-21].

Unlike most assessments, the PT has a repetitive, comprehensive and curriculum-independent nature. The feedback of the PT has an important formative function, aiming to guide and improve students' learning. However, the effectiveness of PT feedback remains uncertain [8, 10-13]. Students often struggle to comprehend or utilise PT feedback [11], and they may lack the agency to translate the feedback into actionable strategies [12]. Agency and its importance in feedback use was also highlighted by Winstone *et al.* [17]. When both PTs and end-of-course tests have summative purposes, students tend to focus on the latter and be less inclined to self-regulate their learning with PT feedback [8]. Furthermore, research shows that students rarely use the PT to reflect and improve their learning [13]. These studies indicate a significant loss of the formative value that PT feedback is intended to provide. Still, acceptance of progress testing is enhanced by sufficient, detailed, personalised, well-timed and specific feedback [4]. However, current literature lacks studies about students' actual use of PT feedback and the factors influencing this feedback use. Therefore, this qualitative study aimed to explore which processes and factors affect medical students' feedback use within a Dutch progress testing context. By exploring this, actions can be undertaken to overcome obstacles, support students' feedback use and improve student feedback literacy.

Methods

Setting

This study was set in a 6-year medicine programme (split into a 3-year preclinical bachelor and 3-year clinical master) at Leiden University Medical Centre (LUMC). In the Netherlands, all medical students are required to participate in a national, curriculum-independent PT four times annually, in addition to their regular course assessments. This results in a total of 24 PT test moments over the course of their study. The PT is a written test consisting of multiple-choice questions (MCQs) that cover all relevant medical disciplines and are stratified into categories. The MCQs include a question mark option that yields no points. Points are deducted if an answer is incorrect [22]. Students receive a 'Good', 'Pass' or a 'Fail' on each PT, and at the end of each academic year, the scores are combined and translated into a summative decision (i.e. study credits). After each PT, the answers are available online, and students receive feedback via e-mail. Students can also access feedback in the online Progress test Feedback System (ProF) [23]. The feedback is presented as an individual score stratified by category and discipline compared with their peer students. This is displayed in the e-mail (*Appendix 1 – Supplemental Table 1*) and graphically in ProF [23]. Students can reflect on the feedback voluntarily with their tutor, who is a medical specialist who offers support throughout the programme. Normally, all students take the PT in a lecture hall under live supervision. During the COVID-19 pandemic, when this study took place, some of the PTs were taken from home via a digital assessment platform. Due to logistic reasons, not all online PTs could be proctored by the online proctoring software. These non-proctored PTs were turned into formative PTs (e.g. no impact on obtaining study credits) because students were able to consult study materials for answering the questions. The online proctored PTs remained summative, just as the PTs taken in the lecture hall.

Study design

Based on the subtle realism paradigm, a qualitative study design was chosen. This paradigm combines a realist ontology with a constructivist epistemology, acknowledging an objective reality independent of our perceptions, while also recognising that our understanding of this reality is shaped by our subjective perspectives [24, 25]. The truth is negotiated through dialogue and the goal of 'objectivity' is an ideal for which to strive. Rigorous research methods, such as purposive sampling, member checking and reflexivity are used to enhance the objectivity and credibility of research findings. Subtle realism also aims to provide a logical and coherent interpretation of the data (i.e. plausibility), instead of proving a definitive cause-and-effect relationship. This paradigm allowed us to provide a nuanced and comprehensive exploration of students' feedback use. We developed an interview guide with open-ended questions about students' feedback behaviour in relation to the PT (*Appendix 2 – Interviewguide*). This study was part of a more comprehensive interview study, in which we also assessed the effects of PT assessment conditions (formative versus summative) on feedback behaviour. After two pilot interviews with fourth-year medical students, individual semi-structured interviews were conducted in April and May 2022.

Sampling and data collection

Medical students were included in the sampling if they participated in at least four out of six PTs in September 2020 to December 2021 and in both formative and summative PT assessment conditions. Maximum variation sampling based on the frequency of ProF use, PT results and study year was used to ensure the representation of multiple perspectives [26]. We first made proportional groups based on

the distribution of ProF logging sessions among all students who participated in the PTs. Within these groups, we aimed to sample an equal number of students from different study years with a 'fail' or 'pass/good'. If a student had failed at least one PT, the student was assigned to the 'fail' group ($n=410$). The other students were assigned to the 'pass/good' group ($n=876$). All student data were derived from the university's student administration system. Participants were invited by e-mail and received an electronic gift card in return for participation. EvW conducted 21 interviews of 30–60 min via online meetings in Microsoft Teams.

Data analysis

The audiotaped interviews were transcribed verbatim and pseudonymized before analysis. Template analysis was used to analyse the interviews, initially after four interviews (when small adjustments were made to the interview guide) and then after all the interviews were completed. Template analysis uses existing literature to formulate a priori themes that guide deductive analysis and the development of successive coding templates [27]. Our a priori themes (*Table 1*) were based on the psychological processes underlying barriers in feedback use as described by Winstone *et al.* [17] and other literature on feedback use [1, 28, 29]. Two independent coders (EvW and FvB) coded Interviews 1–6 in ATLAS.ti 22.0.11.0 for Windows [30]. Afterwards, EvW and FvB discussed the obtained codes together with a third researcher (AL), who coordinates the PT within the LUMC and could therefore provide feedback on the validity and clarity of the initial template from her own expertise. EvW and FvB then coded Interviews 7–14, revised the initial template again and coded Interviews 15–21, after which they determined the final template. This procedure yielded only minor revisions in the initial template. No new themes emerged in the final template, indicating theoretical sufficiency after Interview 14 [31, 32]. The final template (*Table 2*) was discussed with the research team (EvW, FvB and AL). Eventually, EvW reread and recoded all interviews with the final template to ensure all relevant information to answer the research question was included in the template. In this final phase, it became evident that most of the codes did not exist in isolation but co-occurred frequently. EvW coded the relations between these co-occurring codes, discussed these new relations with FvB and leveraged insights from these discussions to develop a thematic map (*Figure 1*). Synthesised Member Checking (SMC) method, which is suggested as an appropriate method within a subtle realism paradigm [33], was performed and yielded the additional theme 'Guidance'.

Table 1. A priori themes used in the template analysis.

A priori themes
1. Awareness
2. Cognizance
3. Agency
4. Volition
5. Score
6. Stakes
7. Time
8. Utility value

Table 2. Final template.

External factors	
1. Stakes	
	Test moment
	Grade
2. Time available	
3. Feedback timing	
4. Feedback message	
	Presentation
	Specificity
5. Guidance	
Internal processes	
1. Awareness	
	Inability to decode feedback
2. Cognizance	
	Poor knowledge of appropriate opportunities
	Poor knowledge of appropriate strategies
3. Agency	
	Sense of disempowerment
	Difficulties with translating feedback into action
4. Volition	
	Lack of interest
	Lack of or change in priority
	<i>Clinical phase</i>
	<i>Other (study) activities</i>
	Grade focus
	Enhancement of self-knowledge
	<i>Performance level</i>
	<i>Future career perspectives</i>

Reflexivity

Researcher reflexivity is required to account for how subjective perspectives shaped the qualitative research process [34]. Our reflexivity statement can be found in *Appendix 3*.

Results

Participant demographics are shown in *Appendix 4 – Supplemental Table 2*. The thematic map illustrates the relations between our themes and codes that describe the processes and factors involved in PT feedback use (*Figure 1*). The findings will be presented as a comprehensive narrative, structured by Winstone *et al.*'s internal psychological processes [17].

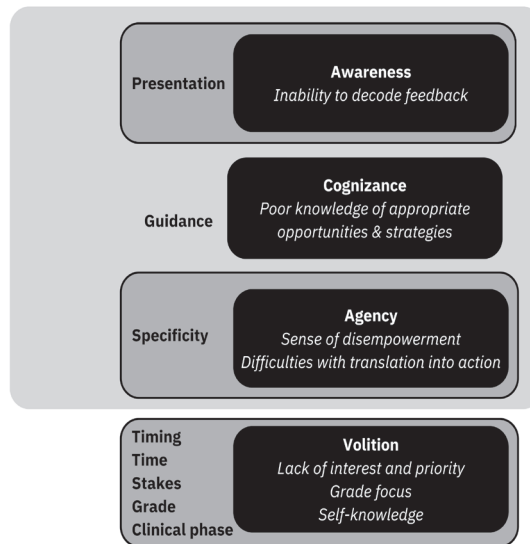


Figure 1. The internal processes (bold white text) with their underlying barriers (italic white text) to feedback recipience are represented as black boxes. The factors that affect the feedback engagement are displayed in the dark grey boxes. Guidance (light grey box) is suggested by students as means to address the challenges related to *awareness*, *cognizance*, and *agency*.

Awareness

Awareness can be conceived as knowing what feedback means and what its purpose is [17]. Almost all students expressed frustration regarding the complexity of the feedback in the e-mail (*Appendix 1 & 2 – Supplemental Table 1 & 2*), which was mainly caused by the unclear **presentation**. As they were **not able to decode this feedback**, the feedback was not used:

“I really never understood that overview in the e-mail. I don’t use it, because I think it is such an abracadabra with all those pluses and minuses.” (Interview #12, Y6)

The majority of students indicated that the feedback in ProF, which is visually presented (also see *Supplemental Figures 1 and 2*), was clearer and therefore easier to understand:

“I find ProF very useful, because you can easily see at a glance in a graph how well you have scored.” (Interview #8, Y3)

Cognizance

Cognizance means that students know about opportunities for seeking support in using feedback and beneficial learning strategies [17]. Several students did not realise they could discuss the feedback with their tutor or that they could access the feedback in ProF. This **poor knowledge of opportunities** prevented them from using the feedback.

However, even students who knew about ProF admitted that they often did not take advantage of it as they were **unaware of which strategies were beneficial** to successfully implement the feedback. Many students lacked knowledge on how to use received feedback:

“My consideration to use the feedback, or prepare for the next PT, depended on how easy it was to study, how broad the topic was. Because, if it’s, for example, you’ve made too many mistakes in internal medicine. Yes, that (topic) is so incredibly broad. I just cannot study that.” (Interview #13, Y6)

Agency

Even when students are cognizant, they still require confidence in the implementation, which is defined as *agency* [17]. Most students failed to **translate the feedback** by e-mail **into action** because of a lack of understanding. Even for students who did understand, or who consulted ProF, the feedback was considered not specific enough making it unclear where to start and how to use it for further learning:

“It (the feedback) is very broad, and there are more questions for each category, and you actually do not know what to do, well there are big differences between the subjects of the questions.” (Interview #1, Y6)

“Yeah, I kind of don’t know where to start. For example psychiatry, I have not scored very well on that lately. Then I think: ‘I have no clue of what I can do. Should I grab the Diagnostic and Statistical Manual of Mental Disorders, or something?’ It would help if we get clear instructions on how to practice, and use the feedback.” (Interview #17, Y5)

Furthermore, many students value the potential of feedback to improve themselves, which gives them a sense of empowerment [17]. However, most students felt **disempowered**, which hindered them from using the feedback, especially when they already passed the PT. This feeling was related to their focus on subject content [35], and direct transferability of knowledge to the next PT rather than appreciating the more long-term development on the different knowledge domains for which the PT is intended.

Volition

To act on feedback, students need to be ‘*ready to engage*’, [36] and thereby show volition: willingness to perform action [17]. Several students were not willing to invest energy in the feedback, because they lacked interest, or gave no priority to the PT. Their perceived late **timing** of the feedback (2 weeks after the PT) reinforced their lack of interest, which suggests that they would rather use the feedback to directly evaluate their performance (score) than to monitor their own growth and learning process. Almost all students mentioned that they had too little **time available** in their (study) schedule to pay attention to the feedback. They usually spent their time on end-of-course assessments, or social activities, which they prioritised over consulting the PT feedback:

“Because we are busy and also have other things, such as examinations, then you think, if this goes well without repetition then yes, you do not want to spend extra time on the feedback.” (Interview #18, Y3)

PT feedback gained priority when the **stakes** (i.e. consequences for study progress) were higher. Stakes were affected by passing, or failing a PT, and by the test moment.

For instance, the last PT of the bachelor or master (Test moment 12 or 24) was high stakes for most students, which motivated students to use previous PT feedback to prepare for the next PT:

"I think that (test moment 24) also played a role, because I think that people are stressed about doing well so they can finish their study. And that is why they also use ProF to look at earlier feedback and be able to prepare well." (Interview #7, Y3)

Also, the **grade** was an important determinant in their volition to act on the feedback. Through this 'grade focus' [17], an insufficient grade incentivised acting on the feedback, whereas a sufficient grade incentivised consulting the feedback superficially, or ignoring it.

Some students changed their attitude towards their study after the transition to the **clinical phase**, which was accompanied by an increasing interest in the feedback:

"I think because I took my studies more seriously in the master compared to the bachelor. In the bachelor, it was more like: 'Pass, fine, check, and move on.' In the master, it was more like: 'Okay, what can I still learn from my mistakes?'" (Interview #20, Y2)

Students who were curious about the feedback mainly used it to enhance their self-knowledge, confirm their overall performance or check their performance on domains of future career interests:

"You're almost at the end of your bachelor's degree, and you're looking at which specialty to choose. At this moment, I really like endocrinology and gynecology, so when I look at the different domains on ProF, I will first have a closer look at the scores on those domains, and ask myself 'Is it justified that I like those specialties?'" (Interview #2, Y3)

Guidance

A lack of guidance limited many students in their feedback use. Students mentioned that they would appreciate more opportunities to practice with PT questions in the curriculum, although there is an existing possibility to practice previous PT questions on the national PT website. They preferred a more detailed explanation of the current possibilities provided by ProF, what the feedback means and how they can use it:

"You could make some practice questions available, so you can see the connection between the PT and the courses and the PT is more integrated in the curriculum. And if, for example, you do not know something, you can directly explore this within the course. I still miss that a bit." (Interview #5, Y3)

Discussion

This study provides an overview of relevant processes and factors that can prevent or stimulate feedback use in medical progress testing and how these processes and factors relate to each other (Figure 1). Most students did not understand the feedback, had poor knowledge of appropriate opportunities and strategies to use the feedback, felt disempowered and insecure about translating the feedback into action, lacked interest and tended to focus on grades.

In contrast to Winstone *et al.* [17], we did not find evidence that students were limited in their feedback use by narrow conceptions of the feedback's purpose. Students seemed well aware that the feedback was meant for self-reflection and could help them grow, in line with the assessment for learning. Moreover, students were willing to consult feedback to learn about their performance in relation to future career perspectives. Students in the clinical phase were also more engaged with the feedback.

Students experienced difficulties in understanding the PT feedback (*awareness*) [15, 17], mainly because the presentation of the feedback in the e-mail was unclear. The visual presentation of the feedback in ProF improves the level of understanding and thereby feedback engagement. In general, students had poor knowledge of the opportunities, such as reflection with the tutor, and strategies to utilise the feedback (*cognizance*) [17], which hindered feedback use. Difficulties with translating the feedback into action (*agency*) [17] were attributed to the lack of specificity of the feedback message, which aligns with prior findings in the PT [11, 12]. In undergraduate medical education, negative perceptions of students also relate to unspecific and unclear feedback [18]. Students reported that they needed more guidance and support to overcome these barriers; they preferred clear and explicit instructions and tools on how to read and handle the feedback. This struggle with appropriate learning strategies and the desire for explicit instructions is also reported in earlier studies [17, 37-39]. Guidance or explicit suggestions on how to address knowledge gaps could enable students to construct appropriate strategies for improvement [19].

Most students were not willing to act upon the feedback (*volition*) [17], primarily due to its perceived insignificance and lack of interest. The delayed timing of the feedback contributed to this lack of interest, indicating a disregard for its importance in enhancing the long-term learning process. The immediate availability of the answer key after the PT was appreciated by the students. Remarkably, several students mistakenly believed that the answers were only available for 24 hours (which actually was the time to appeal for mistakes in the questions and/or answer key), which hindered its use. Several studies show that feedback timing is important, with immediate feedback generally leading to higher task satisfaction [1, 4, 40, 41]. However, most students did not prioritise the PT feedback, and instead, focused on the end-of-course assessments. This aligns with earlier findings where end-of-course assessments were perceived as being higher stakes and more rewarding than the PT [8]. This preference may be linked to the PT's position as a curriculum-independent test and its lack of integration in the curriculum, which is known to be important in promoting student feedback engagement [15, 16, 18-20, 42]. Conversely, higher stakes test moments and insufficient grades on the PT led to increased feedback utilisation. This contrasts with findings of more feedback engagement in high-performing students after objective structured clinical examinations (OSCEs) [43], yet aligns with the concept of '*grade focus*' [17, 44, 45], which has been extensively discussed in the literature as a factor affecting students' receptiveness to and engagement with feedback. Furthermore, the pass/fail decision associated with the PT may contribute to its perception as a purely summative assessment, potentially impeding the intended formative function of the feedback [16, 46].

Feedback engagement was enhanced in students in the clinical phase and in students willing to gain knowledge on their performance and future career perspectives. Although the self-assessment was mainly limited to feedback consultation (a more passive feedback use), this finding provides insights into student incentives to use PT feedback.

Our study demonstrates that the pre-clinical to clinical transition is accompanied by a more serious attitude towards the study and an increased interest in the PT feedback. The feedback can be extra valuable in the clinical phase to address knowledge gaps and apply theoretical knowledge in practice [47, 48].

Strengths and limitations

We positioned our study in the existing literature on feedback use by using a previously established theoretical framework [17] and expanded this framework to elucidate all relevant factors that play a role in the use of PT feedback use. Our sample encompassed a representative sample of medical students from both preclinical and clinical stages with a variety in ProF utilization and PT results. Our findings align with concepts found in earlier feedback research in different educational contexts, encompassing key aspects such as the internal psychological processes [17], specificity [18, 19], feedback timing [1, 4, 40, 41] and grade focus [17, 44, 45]. This convergence suggests that the present findings may also be valid in other settings in which assessment *of* learning is combined with assessment *for* learning, such as a programmatic assessment setting where students are encouraged to use feedback from low-stakes tests to enhance their future learning. Nevertheless, it is crucial to acknowledge that unique features of the PT, such as its' repetitive and comprehensive nature, may affect how students engage with feedback. Additionally, different operationalizations of the PT may affect feedback use as well, as it is shown to influence learner acceptance [4]. For example, other medical schools may put more emphasis on guidance through tutoring, which is shown to stimulate feedback use [42]. Finally, we did not address issues related to Equity, Diversity, and Inclusion (EDI), as these (participant) data were not registered, and therefore not available. Future research could benefit from exploring these issues, as students with specific learning difficulties and from minority groups might face particular challenges with written and graphical feedback.

Implications for practice

The present findings can be used to address and overcome barriers hindering the use of PT feedback, thereby enhancing students' feedback literacy and effective feedback use. These barriers can be overcome by timely feedback provision, clear and specific feedback presentation and guidance to aid students in its effective use. Utilising feedback prompts that describe the learning objectives of the questions could enhance the specificity of the feedback, as shown by Burr *et al.* [49] Even though self-regulation is becoming more important in medical education [50, 51], our study reveals that most students do not use the opportunity to reflect on the feedback with their tutor and may still require explicit instructions and guidance in feedback use. This guidance can be offered by teachers and/or the institution early in the curriculum, as it is known that this shared responsibility is crucial in developing student feedback literacy skills [15, 52]. This may also require a better integration of the PT into the curriculum, as suggested by the interviewed students and earlier studies [42, 46, 53]. Interactive dialogue between students and teachers is important here instead of static information provision [21]. This aligns with key principles of programmatic assessment, in which there is a strong focus on feedback literacy and dialogue. Under such circumstances, students might be more receptive to engage with PT feedback [16, 42].

Conclusion

This study demonstrates that PT feedback use by medical students is hampered by the experience of difficulties related to awareness, cognizance, agency and volition. The stakes, available time, feedback timing, feedback message and lack of guidance contribute to these difficulties and can further prevent feedback use. Student feedback engagement could be enhanced by providing guidance and explicating its relevance for self-regulated learning. Additionally, effective communication and integration of the PT and its feedback within the curriculum can further promote feedback engagement, elevating the PT's significance as a valuable learning instrument.

References

- Hattie J, Timperley H. The Power of Feedback. *Review of Educational Research*. 2007;77(1):81-112.
- Koh LC. Refocusing formative feedback to enhance learning in pre-registration nurse education. *Nurse Education in Practice*. 2008;8(4):223-30.
- Berkhout JJ, Helmich E, Teunissen PW, van der Vleuten CPM, Jaarsma ADC. Context matters when striving to promote active and lifelong learning in medical education. *Medical Education*. 2018;52(1):34-44.
- Dion V, St-Onge C, Bartman I, Touchie C, Pugh D. Written-Based Progress Testing: A Scoping Review. *Academic Medicine*. 2022;97(5):747.
- Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*. 2006;31(2):199-218.
- Black P, William D. Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*. 1998;5(1):7-74.
- Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. *Medical Teacher*. 2011;33(6):478-85.
- Van Berkel HJM, Nuy HJP, Geerligs T. The influence of progress tests and block tests on study behaviour. *Instructional Science*. 1994;22(4):317-33.
- Van Der Vleuten CPM, Verwijnen GM, Wijnen WHFW. Fifteen years of experience with progress testing in a problem-based learning curriculum. *Medical Teacher*. 1996;18(2):103-9.
- Aarts R, Steidel k, Manuel BAF, Driessen EW. Progress testing in resource-poor countries: A case from Mozambique. *Medical Teacher*. 2010;32(6):461-3.
- Given K, Hannigan A, McGrath D. Red, yellow and green: What does it mean? How the progress test informs and supports student progress. *Medical Teacher*. 2016;38(10):1025-32.
- Yielder J, Wearn A, Chen Y, Henning MA, Weller J, Lillis S, et al. A qualitative exploration of student perceptions of the impact of progress tests on learning and emotional wellbeing. *BMC Medical Education*. 2017;17(1):148.
- Schüttelpelz-Brauns K, Karay Y, Arias J, Gehlhar K, Zupanec M. Comparison of the evaluation of formative assessment at two medical faculties with different conditions of undergraduate training, assessment and feedback. *GMS journal for medical education*. 2020;37(4):Doc41.
- Carless D, Boud D. The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*. 2018;43(8):1315-25.
- Molloy E, Boud D, Henderson M. Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education*. 2020;45(4):527-40.
- Watling CJ, Ginsburg S. Assessment, feedback and the alchemy of learning. *Medical Education*. 2019;53(1):76-85.
- Winstone NE, Nash RA, Rowntree J, Parker M. 'It'd be useful, but I wouldn't use it': barriers to university students' feedback seeking and recipience. *Studies in Higher Education*. 2017;42(11):2026-41.
- Cordovani L, Tran C, Wong A, Jack SM, Monteiro S. Undergraduate Learners' Receptiveness to Feedback in Medical Schools: A Scoping Review. *Medical Science Educator*. 2023;33(5):1253-69.
- Lefroy J, Watling C, Teunissen PW, Brand P. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. *Perspectives on Medical Education*. 2015;4(6):284-99.
- Ramani S, Könings KD, Ginsburg S, van der Vleuten CPM. Meaningful feedback through a sociocultural lens. *Medical Teacher*. 2019;41(12):1342-52.
- Ajjawi R, Regehr G. When I say ... feedback. *Medical Education*. 2019;53(7):652-4.
- Lord FM. Formula scoring and number-right scoring. *Journal of Educational Measurement*. 1975;12(1):7-11.
- Tio RA, Schutte B, Meiboom AA, Greidanus J, Dubois EA, Bremers AJA, et al. The progress test of medicine: the Dutch experience. *Perspectives on Medical Education*. 2016;5(1):51-5.
- Maxwell JA, Mittapalli K. Realism as a Stance for Mixed Methods Research. *SAGE Handbook of Mixed Methods in Social & Behavioral Research*: SAGE Publications, Inc.; 2010. p. 145-68.
- Fetters MD, Curry LA, Creswell JW. Achieving Integration in Mixed Methods Designs—Principles and Practices. *Health Services Research*. 2013;48(6 Pt 2):2134-56.
- Onwuegbuzie A, Collins K. A Typology of Mixed Methods Sampling Designs in Social Science Research. *The Qualitative Report*. 2015.
- Brooks J, McCluskey S, Turley E, King N. The Utility of Template Analysis in Qualitative Psychology Research. *Qualitative Research in Psychology*. 2015;12(2):202-22.
- Jonsson A. Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*. 2013;14(1):63-76.
- Harrison CJ, Könings KD, Schuwirth L, Wass V, van der Vleuten C. Barriers to the uptake and use of feedback in the context of summative assessment. *Advances in Health Sciences Education*. 2015;20(1):229-45.
- ATLAS.ti Scientific Software Development GmbH. 22.0.11.0 ed.
- Dey I. Grounding grounded theory: guidelines for qualitative inquiry. San Diego: Academic Press; 1999. 282 p.
- Saunders B, Sim J, Kingstone T, Baker S, Waterfield

- J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity*. 2018;52(4):1893-907.
33. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member Checking: A Tool to Enhance Trustworthiness or Merely a Nod to Validation? *Qualitative Health Research*. 2016;26(13):1802-11.
34. Olmos-Vega FM, Stalmeijer RE, Varpio L, Kahlke R. A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Medical Teacher*. 2022;0(0):1-11.
35. Orsmond P, Merry S. Feedback alignment: Effective and ineffective links between tutors' and students' understanding of coursework feedback. *Assessment & Evaluation in Higher Education*. 2011;36:125-36.
36. Handley K, Price M, Millar J. Beyond 'doing time': investigating the concept of student engagement with feedback. *Oxford Review of Education*. 2011;37(4):543-60.
37. Bing-You RG, Paterson J, Levine MA. Feedback falling on deaf ears: residents' receptivity to feedback tempered by sender credibility. *Medical Teacher*. 1997;19(1):40-4.
38. Winstone NE, Nash RA, Rowntree J, Menezes R. What Do Students Want Most from Written Feedback Information? Distinguishing Necessities from Luxuries Using a Budgeting Methodology. *Assessment & Evaluation in Higher Education*. 2016;41(8):1237-53.
39. Spooner M, Larkin J, Liew SC, Jaafar MH, McConkey S, Pawlikowska T. "Tell me what is 'better'!" How medical students experience feedback, through the lens of self-regulatory learning. *BMC Medical Education*. 2023;23(1):895.
40. Kulik JA, Kulik C-LC. Timing of Feedback and Verbal Learning. *Review of Educational Research*. 1988;58(1):79-97.
41. Bayerlein L. Students' feedback preferences: how do students react to timely and automatically generated assessment feedback? *Assessment & Evaluation in Higher Education*. 2014;39(8):916-31.
42. Heeneman S, Schut S, Donkers J, van der Vleuten C, Muijtjens A. Embedding of the progress test in an assessment program designed according to the principles of programmatic assessment. *Medical Teacher*. 2017;39(1):44-52.
43. Pugh D, Bhanji F, Cole G, Dupre J, Hatala R, Humphrey-Murto S, et al. Do OSCE progress test scores predict performance in a national high-stakes examination? *Medical Education*. 2016;50(3):351-8.
44. Carless D. Differing perceptions in the feedback process. *Studies in Higher Education*. 2006;31(2):219-33.
45. Hounsell D. Towards more sustainable feedback to students. *Rethinking Assessment in Higher Education*. 2007. pp. 101-13.
46. Heeneman S, Oudkerk Pool A, Schuwirth LWT, van der Vleuten CPM, Driessen EW. The impact of programmatic assessment on student learning: theory versus practice. *Medical Education*. 2015;49(5):487-98.
47. Prince KJAH, Boshuizen HPA, van der Vleuten CPM, Scherpbier AJJA. Students' opinions about their preparation for clinical practice. *Medical Education*. 2005;39(7):704-12.
48. Malau-Aduli BS, Roche P, Adu M, Jones K, Alele F, Drovandi A. Perceptions and processes influencing the transition of medical students from pre-clinical to clinical training. *BMC Medical Education*. 2020;20(1):279.
49. Burr SA, Brodier E, Wilkinson S. Delivery and use of individualised feedback in large class medical teaching. *BMC Medical Education*. 2013;13(1):63.
50. van Houten-Schat MA, Berkhout JJ, van Dijk N, Endedijk MD, Jaarsma ADC, Diemers AD. Self-regulated learning in the clinical context: a systematic review. *Medical Education*. 2018;52(10):1008-15.
51. Lucieer SM, Jonker L, Visscher C, Rikers RMJP, Themmen APN. Self-regulated learning and academic performance in medical education. *Medical Teacher*. 2016;38(6):585-93.
52. Carless D, Winstone N. Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*. 2020:1-14.
53. Nouns ZM, Georg W. Progress testing in German speaking countries. *Medical Teacher*. 2010;32(6):467-70.

Appendix

Appendix 1 – Supplemental Table 1

Supplemental Table 1A. Feedback of the progress test with the results per category.

Description categories	Number of questions	Individual					Test moment group (n=57)						
		Correct	Incorrect	?	Score	Correct	Std	Incorrect	Std	?	Std	Score	Std
01 Respiratory system	13	69	31	0	56	68	13	28	12	4	7	57	18
02 Musculoskeletal system	17	59	41++	0--	38	58	11	31	9	11	10	44	14
03 Mental Health Care	16	69	31+	0	58-	75	14	20	12	5	7	68	18
04 Reproductive system	11	45-	55++	0	27--	58	15	29	13	13	13	48	18
05 Blood, lymph, heart and circulation	24	58	25	17+	48	60	13	29	11	11	9	48	17
06 Hormones and metabolism	13	46-	46++	8	-29	57	13	31	14	12	10	46	17
07 Skin and connective tissue	12	83	17	0	78	80	10	17	10	3	6	74	13
08 Personal, social and prevention aspects	17	29--	71++	0--	4--	52	14	38	14	11	10	35	19
09 Digestive system	17	71	29	0--	61	66	12	26	11	8	7	57	15
10 Kidneys and urinary tract	16	69	25	6	59	71	13	21	11	7	8	63	16
11 Nervous system and senses	17	47--	47++	-6	28--	62	13	26	12	12	11	53	16
12 Knowledge about skills	23	48	39	13	33	49	11	40	11	11	9	32	14
Total	196	57-	38++	5-	42-	62	8	29	6	9	6	51	9

-/--/++/+ low respectively high in comparison with the total group. Results are presented in percentages. Std = standard deviation. ? = question mark option use.

Supplemental Table 1B. Feedback of the progress test with the results per discipline.

Description disciplines	Number of questions	Individual					Test moment group (n = 57)						
		Correct	Incorrect	?	Score	Correct	Std	Incorrect	Std	?	Std	Score	Std
Anatomy	12	58	33	8	46	60	15	34	14	6	9	48	20
Biochemistry, molecular and cellular biology and genetics	18	50	44++	6--	34	46	14	31	12	24	14	34	17
Pharmacology	8	62	25	12	54	65	15	27	14	8	9	54	20
Physiology	11	73	27+	0-	62	73	17	18	12	9	12	65	21
Patho-, immuno- en microbiology	10	50	40	10	33-	57	15	34	15	10	10	44	19
Basic-, supportive subjects	59	58	36+	7-	44	58	9	29	7	13	8	47	10
Epidemiology/statistics	7	71+	29	0-	57+	55	23	32	15	12	21	41	26
Metamedica	5	20--	80++	0-	-23--	51	23	38	23	11	14	32	33
Psychiatry/psychology	12	67	33++	0-	54-	73	14	20	12	7	10	65	17
Social medicine	3	33	67+	0-	0	42	26	51	28	8	15	17	37
Behavioural scientific/other subjects	27	56	44++	0-	35-	61	13	30	9	9	10	47	15
Surgery	16	69	31	0-	56	67	13	27	12	6	8	56	17
Dermatology/ENT/ ophthalmology	14	57	36	7	44	63	14	29	14	8	10	53	18
Geriatrics	8	62	38+	0	44	68	17	29	16	3	6	55	23
Obstetrics/Gynaecology	7	43--	57++	0-	21--	60	14	28	17	13	14	49	19
Family medicine	20	40--	55++	5	21--	61	12	34	12	4	5	49	16
Internal medicine	26	73	19	8+	67	73	11	22	9	5	5	64	14
Paediatrics	12	50-	42++	8	32-	60	15	28	13	12	12	48	19
Neurology	7	43	43+	14	19-	50	17	32	17	18	19	37	21
Clinical subjects	110	57-	37++	5	43--	65	8	28	7	7	6	54	10

-/--/+ +/+ low respectively high in comparison with the total group. Results are presented in percentages. Std = standard deviation. ? = question mark option use.

Appendix 2 – Interview guide

Part 1. Own feedback experiences
1. Do you prepare for the progress test?
<i>How do you prepare? What determines whether you prepare for the progress test?</i>
2. Do you consult the result of the progress test?
<i>Which methods do you use to consult the test result? What determines whether you look at the test result?</i>
3. Do you use the result of the progress test?
<i>What do you do with this information? What determines whether you use the feedback?</i>
4. Are you aware of the online feedback system (ProF)?
<i>Why are you not using ProF? What do you think is the reason that you are not aware of ProF?</i>
Part 2. Perception of progress test and feedback
1. What is your perception of the progress test? And which place does it have in your study program?
2. What is your perception of the way(s) the test result is presented to you?
<i>Do you have any suggestions for improvements?</i>

Appendix 3 – Reflexivity statement

The principal and second investigator (EvW, FvB) kept reflective diaries to create awareness of personal expectations, assumptions, and reactions to the participants and data. These diaries were used to guide critical dialogues during data analysis and clarify our interpretations of the data. During data collection, EvW experienced that she could easily relate to the participants, because of her own medical background and experience with the PT. This created an open atmosphere, in which the students felt comfortable to talk openly about their experiences and perceptions. Influenced by her scientific background in (bio) medicine EvW attempted to attain as much objectivity and produce rigorous qualitative research by using maximum variation sampling, member checking, and reflexivity throughout the data collection and analysis. The other researchers were an educational consultant and researcher in medical education (FvB) and a medical doctor with experience in clinical teaching and educational research (AL). Due to his background in cognitive psychology, FvB has been trained to conduct research in an empirical way, based on psychological theories. As such, he supported using theoretical concepts from feedback literature to formulate a priori themes. This theory-driven approach may have influenced the results. AL is a member of the national PT working group and a PT examiner, which might have influenced her perceptions on student behaviour.

Appendix 4 – Supplemental Table 3

Supplemental Table 2. Descriptive characteristics of the participants

Number of logging sessions ^a	0		1		2 to 4		>5		Total year (M/F)	
	Grade	Fail	Pass/ Good	Fail	Pass/ Good	Fail	Pass/ Good	Fail		Pass/ Good
Year 2		11 ^c	3 ^c , 10 ^c		21 ^c		20 ^b , 19 ^c			6 (1/5)
Year 3		7 ^c	18 ^c	4 ^b	9 ^c		5 ^b , 8 ^b	2 ^c	15 ^b	8 (4/4)
Year 5		6 ^c		14 ^c		16 ^c	17 ^b			4 (1/3)
Year 6		13 ^b			1 ^c		12 ^b			3 (2/1)
Total Fail and Pass/Good (M/F)		4 (1/3)	3 (0/3)	2 (1/1)	3 (0/3)	1 (0/1)	6 (5/1)	1 (0/1)	1 (1/0)	

^aProF logging sessions from September 2020 to January 2021. M = male, F = female.

^bMale. ^cFemale.