



# Transparency or Stimulating Meaningfulness and Self-Regulation? A Case Study About a Programmatic Approach to Transparency of Assessment Criteria

Liesbeth K. J. Baartman<sup>1\*</sup> and Frans J. Prins<sup>2</sup>

<sup>1</sup> Research Group Vocational Education, Research Centre for Learning and Innovation, Utrecht University of Applied Sciences, Utrecht, Netherlands, <sup>2</sup> Department of Education, Faculty of Social and Behavioural Sciences, Utrecht University, Utrecht, Netherlands

## OPEN ACCESS

### Edited by:

Bronwen Cowie,  
University of Waikato, New Zealand

### Reviewed by:

Carmen Tomas,  
University of Nottingham,  
United Kingdom  
Jill Willis,  
Queensland University of Technology,  
Australia

### \*Correspondence:

Liesbeth K. J. Baartman  
liesbeth.baartman@hu.nl

### Specialty section:

This article was submitted to  
Assessment, Testing and Applied  
Measurement,  
a section of the journal  
Frontiers in Education

Received: 04 May 2018

Accepted: 12 November 2018

Published: 28 November 2018

### Citation:

Baartman LKJ and Prins FJ (2018)  
Transparency or Stimulating  
Meaningfulness and Self-Regulation?  
A Case Study About a Programmatic  
Approach to Transparency of  
Assessment Criteria.  
Front. Educ. 3:104.  
doi: 10.3389/feduc.2018.00104

This exploratory case study focused on fostering meaning making of assessment criteria and standards at the module level and the course/programme level (the entire study plan), and the role of self-regulation in this meaning making process. The research questions that guided this study are: (1) How can students' meaning making of assessment criteria at the module level be fostered, (2) How can students' meaning making of assessment criteria at the programme level be fostered, and (3) How can self-regulation contribute to students' meaning making process? We explored the design and implementation of a rather new Master's programme in The Netherlands: The Master's Expert Teacher of Vocational Education (METVE). Interviews with three developers, three teachers, and 10 students of the METVE were analyzed. For each research question, several themes were derived from the data. Results indicate that meaning making takes place at the module level by using holistic assessment criteria and evaluative experiences, which allow students to make choices within the boundaries set by the assessment criteria. Meaning making at the programme level is experienced as much more difficult by students as well as teachers. The design of the METVE programme fosters meaning making at the programme level, but METVE teachers also express difficulties supporting this. Finally, we found that students perceive self-regulation as something extra for which they don't have enough time. Self-regulation at the programme level was not explicitly addressed and supported in the METVE, which makes it more difficult for some students to steer their learning process toward the role they are aiming for in professional practice after completing the Master's programme.

**Keywords:** assessment, criteria, transparency, assessment programme, self-regulation

## INTRODUCTION

Higher education aims to build a foundation for professionals in later work settings and social settings. In higher education, the specification of learning outcomes, and standards (the attainment levels) may be desirable in terms of transparency, but an unintended consequence may be to portray to students the idea that learning outcomes are a given (something done to them) and that good work means to work toward criteria set by others (Boud and Falchikov, 2006). In professional

practice, however, no lists or rubrics exist describing what “good work” looks like. Professionals have to be able to form their own complex judgments of their work and that of others, often in collaboration with colleagues, partners, customers, clients, etcetera, in short with all stakeholders directly or indirectly involved in their work (cf., evaluative judgement; Boud et al., 2018). If the above pictures professional practice and what is expected from students in later work settings, what are the implications for assessment and specifically the transparency of assessment criteria? Assessment criteria are often shared with students to communicate expectations and stimulate student performance in the “intended” direction (i.e., most of the times intended by the teacher), mainly at the module level. Transparency of assessment criteria may make clear to students what is expected of them. On the other hand, it may produce students who are more dependent on their teachers and may weaken rather than strengthen the development of self-regulated learning and learner autonomy (Torrance, 2007). From a programmatic perspective, transparency of assessment criteria may prevent students from choosing their own learning goals, their learning tasks and modules and, consequently, prevent them from assembling their own learning path during the curriculum. In other words, transparency of assessment criteria may be detrimental for the development of students’ self-regulatory and lifelong learning skills. Self-regulation refers to self-generated thoughts, feelings, and behaviors that are oriented to attaining goals (Zimmerman, 2000, 2002), which may concern the task level, the module level, the programme level as well as a lifelong learning perspective.

In this contribution, we therefore work out the argument that transparency not necessarily means that students get an exact picture of what is expected of them, but we propose that transparency could instead be viewed as meaning making of assessment criteria, both at the module level and at the programme/curriculum level. We add a curriculum level perspective to the discussion about transparency of assessment criteria, focusing on what is expected of students during the entire curriculum, at the end of the curriculum, and in later working life. In a case study, we explored how students’ long-term development throughout the curriculum can be brought to the forefront and how students’ meaning making of assessment criteria and self-regulatory skills may be stimulated.

The research questions that guided this study are: (1) How can students’ meaning making of assessment criteria at the module level be fostered, (2) How can students’ meaning making of assessment criteria at the programme level be fostered, and (3) How can self-regulation at the module level and programme level contribute to students’ meaning making process? In the remainder of this contribution, we first take the perspective of the module level. Then we shift to the programme level, focusing on students’ long-term learning process toward the programme or graduate learning outcomes. Third, using the framework of Zimmerman (2000; 2002) on self-regulation, we explore the role of self-regulatory skills at the module and programme level in meaning making of assessment criteria. We end with a single exploratory case study to explore meaning-making

of assessment criteria in practice, with varying degrees of success.

## MEANING MAKING OF ASSESSMENT CRITERIA AT THE MODULE LEVEL

In drive for transparency, standards and criteria at the module level (e.g., for assignments or exams) are often made explicit through (long) lists of criteria, benchmarks, rubrics, etc. Several researchers (e.g., Black and Wiliam, 1998; Rust et al., 2003; Wiliam, 2011) argue for the importance of clarifying the intended learning outcomes, because low achievement can be caused by students not knowing or understanding what is expected. On the other hand, students express disappointment about the over-reliance on written criteria to deliver clarity about assessment criteria and the lack of opportunities to internalize standards (Nicol and Macfarlane-Dick, 2006). Assessors use both explicit and tacit knowledge about standards when assessing student work (Bloxham and Campbell, 2010; Price et al., 2011). Bloxham and Campbell (2010) and Hawe and Dixon (2014) showed that when teachers do not share tacitly held criteria with students, this can result in misalignment between the judgments made by students and those made by the teacher. Understanding tacit criteria in a (work) community of practice takes place through an active, shared process rather than a one-way communication of explicit criteria to students. This is also confirmed in a recent literature review on teachers’ formative assessment practices, which also showed the importance of an active role of students in explicating and understanding learning goals and assessment criteria (Gulikers and Baartman, 2017). Students thus need to be actively engaged to develop a conceptualization of what constitutes quality if they are to improve their work and reach higher levels of performance (Sadler, 2009).

The ability to assess a piece of work against contextually appropriate standards is at the heart of “evaluative judgment” (e.g., Boud et al., 2018; Panadero and Broadbent, 2018). Research into evaluative judgment also offers suggestions for pedagogical practices in the classroom to stimulate students’ evaluative judgment capacity, which fits nicely with our ideas about meaning making of assessment criteria. Panadero and Broadbent argue for the importance of peer assessment and self-assessment, as these activities enhance evaluative judgment capacity. Other strategies include the use of scaffolding tasks, rubrics and exemplars to clarify and discuss assessment criteria and expectations (e.g., Fluckiger et al., 2010; Conway, 2011). Students can be confronted with a wide variety of authentic works, from other students attempting the same task and/or authentic products from “the real world,” review these good and bad examples and distill success criteria together (Fluckiger et al., 2010; Willis, 2011; Hawe and Dixon, 2014). In higher education, Fluckiger et al. (2010) describe and evaluate four strategies aimed to involve students as partners in the assessment process, to develop a learning climate, and to help students use assessment results to change their learning tactics.

Altogether, in a meaning making process these activities stimulate students to discover that different responses to an

assessment task may all result in valid products that comply with the quality criteria fit for the task. Or as Conway (2011) explains about his history lessons: “if the success criteria are shared and the students understand both what they are working toward and why, then they can take a lot of responsibility and we can allow a lot of variety. (p.4).”

## MEANING MAKING OF ASSESSMENT CRITERIA AT THE PROGRAMME LEVEL

So far, our discussion focused on transparency and meaning making of assessment criteria at the module level, for single assignments or exams. However, the ultimate goal of curricula in higher education is to prepare students for working and social life, and lifelong learning. Assessment involves making judgments about quality and identifying appropriate standards and criteria for the task at hand. This is as necessary to lifelong learning as it is to any formal educational experience (Boud, 2000). What constitutes quality is not a matter of one specific assignment or piece of work, but we view quality as a generalized attribute that can take specific forms or meanings in different contexts. Higher education aims to prepare learners to undertake such judgmental activities and to identify whether their work meets whatever standards are appropriate for the task at hand. To do so, Bok et al. (2013a) and Bok et al. (2013b) focus on stimulating students' feedback-seeking behavior during an entire assessment programme. Students seek feedback from various sources during their clinical clerkships, depending on personal and interpersonal factors such as the students' goal-orientation (focused on learning or on keeping a positive self-image) and the anticipated costs and benefits of the feedback. This feedback seeking behavior and judgments of quality also authentically mirror the ways many quality appraisals are made in everyday and work contexts by professionals.

Consequently, when it comes to transparency and meaning making of assessment criteria, this is not only important at the module level, but also (and maybe even more) at the programme level. In the Netherlands, the context of this study, we observe a drive toward detailed module specifications and explicit assessment criteria, a development Hughes et al. (2015) and Jessop and Tomas (2017) also describe in the UK. In this contribution, we therefore add a programme-level perspective to the discussion about transparency. In programme-focused assessment (van der Vleuten et al., 2012; Bok et al., 2013a) an arrangement of assessment methods is deliberately designed across the entire curriculum, combined and planned to optimize both robust decisions about students (summative) and student learning (formative). Rather than focus on specific or isolated assessments at the module level, a programme perspective focuses on the holistic developmental goals of the programme as a whole (Rust et al., 2012; Hughes et al., 2015). It follows then that such assessment is integrative in nature, trying to bring together “data points” or sources of information about students' development that represent—in varying ways—the key programme outcomes (PASS position paper, 2012). Formatively, assessment activities are viewed as information sources that

provide a constant and longitudinal flow of information about student learning (Heeneman et al., 2015). The balance is shifted from summative to formative assessment to encourage students to think about longer term development rather than short term grade acquisition (Heeneman et al., 2015). An important starting point for programme-focused assessment is an overarching structure: the specification of the programme or graduate learning outcomes (Lokhoff et al., 2010; Hartley and Whitfield, 2011) and a number of levels or stepping stones that describe the development process toward these programme learning outcomes. These stepping stones are comparable to the learning progressions mentioned in Gulikers and Baartman's review (2017) on teachers' formative assessment practices. Key to stepping stones or learning progressions is that these specifications enable teachers and students to monitor progress on a longer term.

Programme learning outcomes are necessarily described in a holistic way as they need to capture the diversity of the (future) professional work context. Some concepts—like what constitutes a “good” or “tasty” dish—are in principle beyond the reach of formal definition. Experts in a professional domain can give valid and elaborate descriptions of what quality looks like in a particular specific instance (e.g., a cook can distinguish a good from a bad dish), but they are unable to do so for general cases. Sadler (2009) therefore argues for the use of holistic assessment criteria, because students need to be induced into judging what quality entails, without being bound by tightly specified criteria. Analytic grading constrains the scope of student work (one solution) and offers little imperative to explore alternative ways forward. The discussion between holistic and analytic or task-specific criteria is a complicated one, especially when it comes to meaning making of assessment criteria at the programme level. Previous research shows the advantages of task-specific criteria (Weigle, 2002; Jonsson and Svingby, 2007), as these criteria provide clear directions to students about what is expected. Govaerts et al. (2005) provide a more nuanced picture indicating that starting students prefer more analytic criteria, whereas experienced students prefer holistic criteria. As programme-focused assessment aims to encourage students to focus on long-term learning processes instead of short-term grade acquisition (Heeneman et al., 2015), task-specific criteria might be less suitable as these criteria tend to focus students on the task at hand, and less on what the student's performance on this specific task tells about the student's long-term development.

A programme perspective on transparency and meaning making of assessment criteria is helpful, because students should not be considered competent at judging the quality of their own and each other's work from the start of the curriculum. Moreover, it is not realistic to expect students to become expert judges of their own work within the scope of a single module. But as the programme proceeds the students' judgments of their own work should gradually reflect the (broad, holistic) programme learning outcomes and expectations of working professionals. The design of an assessment programme should give students insight in their learning and longitudinal development, ensure the main focus is on meaningful feedback to enable students to develop toward the programme learning outcomes (Bok et al., 2013a). If students

are to make meaning out of programme learning outcomes, then processes of feed up, feedback, and feedforward require a dialogue between students and teachers or between students and peers (Nicol and Macfarlane-Dick, 2006). Boud (2000) therefore argues that assessment should move away from the exclusive domain of the teacher/assessor into the hands of learners. Peer assessment could be implemented purposefully at different stages of the assessment programme. When they are still learning to become expert judges of quality, students need structure, and guidance when assessing their peers' work. Altogether, a programme perspective to transparency and meaning making of assessment criteria shows how students' meaning making process need to be purposefully guided over the period of the entire curriculum, and (formative) assessment moments, and evaluative judgment experiences should be purposefully planned and used.

## TRANSPARENCY, MEANING MAKING, AND SELF-REGULATORY SKILLS

Assessment and self-regulatory skills are intertwined in different ways. For instance, as Wiliam (2011) argues, an important aspect of formative assessment is activating learners as the owners of their learning process. In the same vein, Brown and Glover (2006) identified three levels of feedback: that which provides information about a performance; that which provides explanation of expected standards; and that which enables learners to self-regulate future performances. Also, Clark (2012) specifically links formative feedback to self-regulation, indicating that the objective of formative feedback is to support self-regulated learning and give the learner the power to steer one's own learning (p. 210). Furthermore, the use of specific assessment instruments may also have impact on students' self-regulation. As an example, Panadero and Romero (2014) examined the effects of using rubrics on students' self-regulation and concluded that "it is probable that the use of rubrics has a considerable impact on self-regulation, as its use promotes the strategies that have been shown to have the biggest effect on self-regulation interventions: planning, monitoring and evaluation" (p. 141). In other words, assessment and specific assessment instruments may foster students' self-regulation. Zimmerman (e.g., 2000) distinguishes three cyclical phases of self-regulation, that is, the forethought phase (occurs before efforts to learn), the performance phase (occurs during behavioral implementation), and the self-reflection phase (occurs after each learning effort). Especially in the first and third phase, assessment and assessment criteria may play a significant role. In the forethought phase, important processes are goal setting and outcome expectations. Even though very explicit and analytic assessment criteria may make clear to students what is expected of them, it may also produce students who are more dependent on their teachers and may weaken rather than strengthen the development of learner autonomy (Torrance, 2007). Autonomy may be understood as the ability to take care of one's own learning (Panadero and Broadbent, 2018). Consequently, transparency of assessment criteria may be detrimental for the development of students' self-regulatory skills because it may prevent them from

choosing their own learning goals and assembling their own learning path. In the self-reflection phase, self-evaluation (i.e., self-assessment) and causal attributions are main processes, and may be based on the same assessment criteria and standards. When we zoom in on meaning making of assessment criteria, meaningful assessment criteria will probably make it easier for students to formulate personal learning goals in the forethought phase of self-regulation as well as to self-evaluate after the learning efforts. Thus, we argue that meaningful assessment criteria may challenge students to regulate their own learning and increase their autonomy. Furthermore, a high level of self-regulatory skills may have impact on the way students deal with and interpret assessment criteria and standards. Panadero and Broadbent (2018, p. 82) argue that students who know how to self-regulate and to judge their own work can be autonomous and have more opportunities to develop evaluative judgement capacities (i.e., the ability to assess the quality of a piece of work).

So far, we addressed the relation between assessment (including transparency and meaning making) and self-regulatory skills mainly at the module level. But also at the programme level assessment and self-regulatory skills have a reciprocal relation. Zimmerman (2002) argues that self-regulation is important because a major function of education is the development of lifelong learning skills. Lifelong learning skills are necessary during an educational programme but also afterwards, in professional life. Assessment criteria and standards concerning the programme learning outcomes (as well as specific modules) should be meaningful in the sense that students should be able to grasp what these learning outcomes (i.e., assessment criteria and standards at the programme level) may mean for their own learning path and their professional development. Students' meaning making process of assessment criteria at the programme level may contribute to their ability to make choices and to become the professional they are aiming for. One of the goals of an entire curriculum and assessment programme can be to foster students' self-regulatory skills. An assessment programme hardly fosters these skills if the teachers tell students what to do and what to aim for. Instead, an assessment programme should reward students for identifying gaps in their abilities and developing effective ways to correct those gaps (Dannefer and Henson, 2007). In other words, an assessment programme that allows students to set their own learning goals and to formulate their personal assessment criteria and standards more explicitly calls for self-regulatory skills. Only a very small number of studies in a review on teachers' formative assessment practices (Gulikens and Baartman, 2017) showed examples of teachers allowing their students to set their own learning goals or allow students' learning goals to develop and change throughout a course or longer learning trajectory (e.g., Parr and Limbrick, 2010; Kearney, 2013; Lorente and Kirk, 2013; Hawe and Dixon, 2014). Concluding, we argue that an assessment programme should activate learners as the owner of their own learning process and allow learners to create meaningfulness of assessment criteria and standards by combining their own learning goals with available holistic assessment criteria and standards.

Academic year 1		Academic year 2		Academic year 3	
Semester 1	Semester 2	Semester 3	Semester 4	Semester 5	Semester 6
<b>Professional Theory and Impact 1</b> Key task 1 t/m7 5 EC		<b>Professional Theory and Impact 2</b> Key task 1 t/m7 5 EC		<b>Professional Theory and Impact 3</b> Key task 1 t/m7 5 EC	
<b>Teaching in vocational and professional education</b> Key task 1, 4, 5 10 EC	<b>Perspectives on vocational and professional education 1</b> Key task 1, 2, 3, 4, 5 10 EC	<b>Assessing in vocational and professional education</b> Key task 2, 4, 5 10 EC	<b>Designing vocational and professional education</b> Key task 3, 4, 5 10 EC	<b>Perspectives on vocational and professional education 2</b> Key task 5 10 EC	<b>Masterthesis</b> Key task 6 10 EC
<b>Practice-based research 1</b> keytask 6 5 EC		<b>Practice-based research 2</b> keytask 6 5 EC		<b>Practice-based research 3</b> keytask 6 5 EC	

FIGURE 1 | Master's Expert Teacher of Vocational Education Curriculum.

## METHODOLOGY

This study can be characterized as an in-depth single case study (Yin, 2014), which serves to empirically explore how designers, teachers, and students experience meaning making of assessment criteria and the role of self-regulation. In this single case study, we explored the design and implementation of a Master's programme in The Netherlands: the Master's Expert Teacher of Vocational Education (METVE). The METVE is a relatively new Master's programme and currently running for the third year (including a pilot year with 5 students). The design of this Master's programme (partly) includes the arguments about meaningfulness and self-regulation at the module level and programme level discussed above. Therefore, the METVE provided an interesting case to explore meaning making of assessment criteria in practice. The aim was to explore how the transparency of the programme learning outcomes was perceived and to reveal advantages and disadvantages of implementing transparency in terms of creating meaningfulness and fostering self-regulatory skills at the module level and at the programme level.

### Context of the Study (METVE)

The METVE is a part-time Master's programme for teachers of vocational subjects working in preparatory secondary vocational education (VMBO), senior secondary vocational education (MBO) and higher vocational education (HBO). In order to enhance the quality of vocational education in the Netherlands, the Educational Council of the Netherlands (2013) recommended increasing the standards of teachers: 25% of teachers of vocational subjects must have a Master's degree.

For higher vocational education, the aim is that by 2020, 100% of all teachers have at least a Master's degree. The METVE was developed to reach this goal and started in 2015 with a small pilot group of 5 students, continuing in 2016 and 2017 (15 and 20 students, respectively). All METVE students have (many) years of working experience as a teacher in vocational education, a bachelor degree (or equivalent) in their own occupational field (e.g., nursing, business, engineering) and a teaching certificate. Working in vocational education themselves, the METVE students have varied experiences when it comes to assessing their own students. In vocational education, students are generally assessed using a combination of knowledge tests, practical demonstrations, and assessments in the workplace. Competence-based standards are determined at the national level for the different occupations (for a more elaborate explanation about assessment in Dutch vocational education, see Baartman and Gulikers, 2017). The curriculum design of the METVE can be characterized as follows. The METVE curriculum works toward seven core tasks of vocational teachers, which are defined as the programme goals (or attainment levels): guiding students in vocational education, assessing students in vocational education, designing learning environments, connecting learning in-school, and outside school settings, connecting subject knowledge with the profession, doing practice-based research and professional development as a teacher. The METVE curriculum is divided into 5 or 10 ECTS modules in which students work on authentic assignments representing one or more of the seven core tasks. **Figure 1** gives an impression of the various METVE modules and core tasks.

For the entire METVE curriculum a general rubric has been developed based on the Dublin Descriptors, developed

as part of the Bologna Declaration and the Framework for Qualifications of the European Higher Education Area. The Dublin Descriptors provide descriptions of the different levels of higher education, developed to improve transparency and comparability of qualifications across Europe. For the METVE curriculum, the level descriptions of Bachelor (the entry level) and Master (the intended end level) were used. The Dublin Descriptors refer to the following five dimensions: knowledge and understanding, applying knowledge and understanding, making judgements, communication and learning skills. The general rubric serves to monitor and guide METVE students' long-term development process from bachelor-level toward master-level, across the different modules of the METVE. An English translation of the general rubric can be found in **Table 1**. For the different modules of the METVE, this rubric has been specified or contextualized into assessment criteria—again in a holistic way—to represent the assignment of that module, and what for example bachelor-plus performance looks like for that assignment. Within all modules, METVE students are assessed on the core tasks which are central stage in the assessment criteria. For example, in the first module, the METVE students develop a lesson plan for their own students in which they strive to connect learning across different sites inside, and outside the school. METVE students are assessed on three core tasks based on their lesson plan: guiding their vocational students, connecting learning inside and outside school and connecting subject knowledge with the profession. METVE students do not receive a score for their assignment (the lesson plan), but three separate scores for the core tasks. This way, the core tasks are in plain sight when assessing METVE students and students and teachers can monitor METVE students development on the core tasks throughout the curriculum (cf. van der Vleuten et al., 2012).

## Participants

Interviews were carried out with three developers/teachers, three teachers, six 1st-year students and four 2nd-year students. The three developers were involved in the design process of the METVE and the discussions about the underlying rationale of the master's programme. They were interviewed both in their role as developer and in their role as teacher. The three teachers got involved in the METVE in a later stage and had one to 2 years of experience in teaching in the METVE. Teachers and students participated voluntarily and signed a consent form for their participation.

## Interviews

In-depth (group) interviews were used to explore the self-reported experiences of METVE students and teachers with regard to the assessment criteria and the experienced meaning-making and self-regulatory activities. The individual interviews with the developers/teachers lasted 1 hour. One teacher was interviewed individually (30 min), and two teachers were interviewed together (60 min), based on possibilities in their teaching schedules. The students were interviewed in two group interviews, one for the 1st-year students and one for the 2nd-year students. Student interviews lasted 1 hour and were conducted using Adobe Connect (virtual classroom), a digital system the

students were familiar with in their webinars. All interviewed were audiotaped and transcribed verbatim. The interviews were carried out by the first and second author together. Because the first author is one of the developers/teachers of the METVE, the second author took the lead in the interviews to guarantee independence and stimulate the participants to freely express their opinion. Interview questions were asked about three topics: (1) transparency and meaningfulness of assessment criteria for the different modules / assignments of the MEB, (2) transparency and meaningfulness of the entire assessment programme, and (3) how self-regulation and ownership are addressed and experienced. Examples of questions asked to the teachers are: "how do you work with the assessment criteria in your module" and "how do you experience the connections between the modules in terms of students' long term development"? Examples of questions asked to students are: "how do you experience the freedom of choice when it comes to the assignments" and "what do you do to get an idea of what is expected of you in an assignment"? Besides the interviews, documents about the METVE were collected, such as policy documents, course guides and assessment forms. Some developers/teachers referred to these documents in their interviews and provided digital versions of the documents after the interview.

## Analyses

Thematic data analysis was carried out in three rounds by the two authors collaboratively. Template analysis was used, which consists of a succession of coding templates and hierarchically structured themes that are applied to the data (Brooks et al., 2015). After the interviews with the first and second developer/teacher, the 1st year and 2nd year students had been carried out, a first version of the template was developed by the two authors collaboratively, based on their experiences during the first interviews. The three research questions were used as an analytic framework: separate thematic codes were developed for each of the research questions. All fragments were coded using Excel: each fragment could be assigned a theme for either one, two, or all three research questions using pull down menus. This way, some fragments could be assigned to multiple themes when applicable (in practice, fragments never applied to all three research questions).

This first version of the template was used to analyze the interview with developer 1. Both authors coded the interview independently and their analyses were discussed in a meeting, resulting in adaptations to the themes (e.g., definitions were sharpened and themes were added that emerged from the data). The meeting resulted in a second version of the template. Also, the interviews with the other three teachers and developer 3 were carried out, in which the researchers asked follow up questions on themes that had not become clear in the first round of the analysis. Using the second version of the template, both authors independently coded the 1st and 2nd year student interview and (again) the interview with developer 1. The analyses were again discussed in a meeting, resulting in only minor changes in the themes (sharpening definitions so all fragments fitted the description of the themes). This resulted in the third and

**TABLE 1** | General rubric of the METVE (translated from Dutch).

Dublin descriptor	Bachelor	Bachelor +	Master	Master + (excellent)
Knowledge and understanding	Student has knowledge and understanding based on specialized handbooks. Describes theories separately.	Student has knowledge and understanding, partly based on scientific literature. Compares different theories.	Student has deep knowledge and understanding based on scientific literature. Connects theories leading to new (practical insights).	Student has profound knowledge and understanding based on a large body of scientific literature. Connects theories leading to new (scientific) theories.
Applying knowledge and understanding	Student applies knowledge in simple educational problems. Applies a theories to his/her own educational context.	Student applies knowledge in complex educational problems. Chooses theories that fit the own educational context.	Student applies knowledge in complex, new and/or interdisciplinary educational problems. Chooses and argues the best fitting theories for the own educational context, based on a comparison of scientific literature.	Student applies knowledge in complex, new, interdisciplinary educational problems that go beyond the own educational context. Chooses and argues the best fitting combination of theories based on multiple scientific sources and the own experience.
Making judgments	Student analyses educational problems in a partly systematic way. Critically investigates the own educational context. Identifies strengths and weaknesses based on single theories and practical experience.	Student analyses educational problems in a mostly systematic way. Critically investigates the own educational and broader social context. Identifies strengths and weaknesses and makes informed judgments based on theories and practical experience.	Student analyses educational problems in systematic way. Critically investigates the own educational and broader social context, in relation to scientific theories. Identifies strengths and weaknesses and makes informed judgments based on (im)possibilities in practice and research.	Student makes a thorough and critical analysis of educational problems using multiple perspectives. Critically investigates the own educational and broader social context, comparing practical (im)possibilities and scientific theories. Choices are made based on scientific theories and validated in practice.
Communication	Student participates in the own educational institute. Communicates ideas and possible solutions with colleagues.	Student participates in a network in the own educational institute and surrounding industry. Discusses ideas and possible solutions with colleagues and representatives from industry partners.	Student participates in a network in the broad educational and industrial context. Discusses ideas and possible solutions, including underlying motives and considerations with colleagues, and representatives from industry partners.	Student participates in a network in the broad educational and industrial context. Takes initiatives and shares knowledge, for example in publications, presentations and workshops. Discusses ideas and possible solutions, including underlying motives and considerations with colleagues and representatives from industry partners.
Learning skills	Student shows continuous development by means of reflection on his/her own professional performance in the broad context of vocational education.	Student shows continuous development by comparing his/her own professional performance to theory (doing and thinking). Analyses the discrepancies between doing and thinking.	Student shows continuous development with a high degree of autonomy. Learns by means of conceptualizing and concretizing.	Student shows continuous development with a high degree of autonomy. Learns by means of conceptualizing and concretizing. Takes the initiative to learn from complex situations.

final version of the template, which was used by both authors independently to (re)analyze all interviews (see **Table 2** for the final template). Finally, both authors independently selected all fragments belonging to a theme (using the pull-down selection menu), re-read the fragments and made a summary of the theme together with some illustrating examples. This was done for all themes separately. In a meeting, the summaries of the themes belonging to research question 1 were discussed (4 themes). The authors read each other's summaries and made notes when they noticed (big) differences. The main question that guided the discussion was: do we see any results/conclusions that do not logically emerge from the data? Some differences between the authors did appear, mainly because of overlap between some of the themes. For the summaries for research questions 2 and 3 both authors again audited each other's summaries and made notes of differences they encountered. The first author used the discussion and the notes to make a final summary per theme, which was checked by the second author.

## RESULTS

The results are presented per research question and fragments are added as illustrations of the themes that appeared from the data.

### Fostering Meaning Making of Assessment Criteria at the Module Level

For the first research question "How can students' meaning making at the module level be fostered," four themes were derived from the data, related to (1) holistic assessment criteria, (2) meaning making by means of practical relevance, (3) meaning

making by evaluative experiences, and (4) making connections between the general rubric and assignments at the module level.

The interviews showed that in the design of the METVE programme, a holistic approach to assessment was deliberately chosen because of the diversity of the student population, who all work in different domains and levels of vocational education. Also, one developer explained how holistic criteria do more justice to the complexity of tasks students encounter in practice: "an educational argument is that we think if you take a more holistic view, well that is actually how the core tasks, how complex they are. So there is no recipe for carrying out vocational education. That recipe does not exist. So we cannot work out the assessment criteria from A to Z. Because they just not exist. So you have to assess holistically" (developer/teacher nr1).

Though the teachers seem to value the holistic assessment criteria, METVE students reacted in a more diverse way: "well, it may depend on me as a person ... I ... I think this broader framework and the fact we are not pushed into a certain direction, it also gives you the possibility to work out an assignment in your own way" (2nd year student nr1) or "I notice, but as a person I work in the technical domain ... and there you are pragmatic, I like to have a guiding principle to deliver something" (2nd year student nr2). METVE students sometimes seem to feel uncertain about what is expected of them. As one of the students described it: "You have freedom in how to do the assignments, but this freedom can also make you insecure, because you can't exactly pin down what the purpose is of what you have to show" (2nd year student nr4). Both teachers and students still need to build up impressions of what good work might look like in all its possible varieties. One developer explained: "But students also have to build up these images, but we as teachers also need to build up the images, I experience" (developer/teacher nr2). This student agrees with the holistic assessment criteria, but also expresses his need for certainty: "But I think, when you use a holistic assessment model—I understand the goal of that very well—that you also need a further explanation of the module, or the goals of the module, from the start. So if you ask me, it is connected: you either get your information from the assessment criteria or it needs to be made clear that the assessment criteria do not contain all specific information" (2nd year student nr 3).

The diverse METVE student population and holistic assessment criteria bring us to the second theme, namely the increase of meaningfulness by contextualizing assignments to the METVE students' work context. It is this connection to their own work field that makes the assignments relevant or meaningful, and indirectly, also increases the meaningfulness of the holistic assessment criteria. The module assignments explicitly require METVE students to explore developments in their own work context, for example by talking to colleagues, managers, and experts. METVE students thus address a relevant problem or question experienced in their own work context and they try to realize impact on their work context, for example by the products they develop: "we could make our own choice when it comes to the content of the assignment. I chose self-management, because I work with elderly clients. Yes, it was very meaningful for me.

**TABLE 2** | Final template used for data analysis.

Module level	<ul style="list-style-type: none"> <li>Student and teacher experiences of holistic assessment criteria</li> <li>Meaning making by means of practical relevance</li> <li>Meaning making by organizing evaluative experiences</li> <li>Making connections between the general METVE rubric and assignments at the module level</li> </ul>
Programme level	<ul style="list-style-type: none"> <li>Design of the METVE curriculum</li> <li>Teacher and students activities to support meaning making of assessment criteria at the programme level</li> <li>Students get an impression of their development throughout the programme</li> <li>Conditions or prerequisites to achieve meaning making at the programme level</li> </ul>
Self-regulated learning	<ul style="list-style-type: none"> <li>Students' own development at the module level</li> <li>Students' own development at the programme level</li> <li>Experiences of lack of time</li> <li>Students' professional role after graduation</li> <li>Making self-regulation more explicit</li> <li>Teachers' support of self-regulation</li> </ul>



I really experienced an added value ... I could also ask better questions to my pupils" (1st year student nr1). And another student: "You explore what is going on in your department, what relevant issues are, what you like to know more about, in collaboration with your team. So I discuss with my team which research question I address. And also in the module about assessment, I discussed what issues there are, what we like to have an answer for" (2nd year student nr2). This meaning making process goes two ways, from the METVE assignment to the work context of the METVE student and the other way around. Students thus move within the boundaries set by the assessment criteria and within these boundaries experience freedom to contextualize the assignments: "our choice is the topics ... I think, with all assignments, within the boundaries you are free to make choices" (2nd year student nr1). A METVE developer explained how they safeguard the boundaries the students move within, so choices METVE students make fit within the holistic assessment criteria: "we do tell them to go to their own school, talk to colleagues, managers, what is interesting. Sometimes they get a bit stuck, like my manager wants this, but I don't know if it is relevant to the METVE. So that is a step we safeguard, is it a relevant assignment" (developer/teacher nr2).

Themes 3 and 4 at the module level are related to each other and portray how teachers and students work on meaning making by means of evaluative experiences (theme 3). The goal of these evaluative experiences is to make connections between the assignments the students are working on, the assessment criteria and the general rubric (theme 4). Some evaluative experiences were explicitly designed in the curriculum, for example the formative moments during the modules in order to give feedback while students are still in the midst of the meaning making process and can still make choices and adaptation in their assignments. One of the developers explained: "well, for all modules ... it is a formative process. Student work from moment zero toward the end result that will be assessed. So it is not just some separate small assignments, that you first get assignment 1, and then 3, and then 6 and all small assignments together result in ... no, they actually work all the time toward that end product. So all feedback they get is about the end product" (developer/teacher nr1). Or: "it is an holistic assessment about the three [formative part-assignments], in which I explicitly give the message to students that the part assignments have a formative goal, and that I give feedback to part assignment 1 based on the assessment criteria and the Dublin Descriptors so they can show growth within my module" (developer/teacher nr3). Other evaluative experiences were not explicitly designed and depend on individual METVE teachers. Examples of evaluative experiences used in the lessons are: discussions of student assignments using the assessment criteria, peer feedback activities, peer group intervision, teacher feedback, and modeling how you assess as a teacher. These quotes show how METVE teachers and students tell about the evaluative experiences:

"I think it is very worthwhile, to do it often. And it does happen in some instances. In module XXX the teacher projected a student's piece of work on the smart board, and well, how you would assess

it based on the assessment criteria. We could do that more often, it would help immensely" (1st year student nr4)

"Some time ago we discussed the rubrics during the XXX lessons, because we are going to do peer assessments ... and actually we worked out the rubric in pairs. That was very clarifying because you actually, because per theme [assessment criterion] you discuss well, how you would assess someone as a peer assessor. I found that a very interesting addition, to make the holistic more concrete" (2nd year student nr2).

"I used some activities in the lessons, I let them compare some examples. They had to bring their own assignment and in small groups, using the assessment criteria—not really the rubric—they looked for good examples of what assessment criteria might look like. And they made big posters of the examples" (developer/teacher nr3).

METVE students express the value of the evaluative experiences, but they would like to have more evaluative experiences during the lessons, and especially at the beginning of a module to get an impression of what is expected: "I would appreciate it very much if it were at the beginning of the module. And I would like to discuss it in class. Like look, this is the level you are at now, and this is what you are heading for, the ultimate goal is master level" (1st year student nr2). There is a demand to explicitly discuss the assessment criteria, because the general rubric contains concepts that students find hard to grasp: "well, as a student you apparently like to be taken by the hand a bit, so in the lessons you like the teachers to help you, in the right direction ... and then you assume it is all right. And the assessment criteria I think, I see a number of sentences, but what is exactly meant by them?" (2nd year student nr4). The students also expressed the limits of peer feedback: "it is also about that you have to know how to interpret the assessment form. Because when you have not made anything yet, and if you do not know how a teacher would assess it, then it is difficult to grasp. Because we also gave each other feedback and one of us used the assessment form and the other three did not. You have to learn to read it" (1st year student nr6).

The goal of the evaluative experiences, but also the design of the METVE programme, is to make connections between student assignments, the assessment criteria and the general rubric. These connections were meant to increase meaningfulness to METVE students: they can analyze how their specific assignment, choices, and work context relate to the assessment criteria and thus whether they comply with the assessment criteria and the required (master) level. In practice, METVE students do use the assessment criteria to find out what is expected of them at the module level, for example by reading the module guide, and assessment criteria and comparing them to their own work. Students also expressed they rely on the teacher: "I think I do not use it [the assessment form] very often, less than I should ... I think I just use the lessons to know what is meant by the assignments, and then I just get to work. And actually, I just use it at the end as a kind of checklist to check whether I did everything that is expected. But during the lessons you get so much feedback and input from the teachers, I actually lean more from that than from the assessment form itself" (2nd year

student nr4). In the design of the METVE, the general rubric has been translated in assessment criteria for the different modules. These assessment criteria are more specific and concrete and are thus more (directly) meaningful to students. For students, the relationship between their assignments and the general rubric is not always clear. One teacher told: “well, they do read the rubric ... or I think so ... but in the end they look more at the part assignment and the criteria for the assignments. Even if we made an explicit connection between the part assignment and how they are linked to the rubric. But they do not really look at that ... well, it is more contextualized ... it is less general” (teacher nr3).

## Fostering Meaning Making at the Programme Level

Four themes illustrate how meaning making of assessment criteria is fostered at the programme level: (1) by the design of the METVE programme, (2) by teacher and student activities, (3) because students get an impression of their development throughout the programme, and (4) conditions to be met if meaning making is to take place at the programme level.

First, meaning making at the programme level is fostered by the design of the METVE programme. A programmatic approach to assessment is used (van der Vleuten et al., 2012) within constraints such as the demand for a modular curriculum: “Or course, we had the idea of gradual development in the curriculum, a development line. That was the first dilemma in our curriculum, because we wanted a nice progression and an increase in complexity, while actually the demand was that students should be able to do separate modules. That you do only one module. So that is kind of tension in our curriculum” (developer/teacher nr1). Important elements of the METVE design that foster students’ meaning making and long-term learning processes are the general rubric to assess student work in the various modules and the fact that students can show growth on the core tasks throughout the curriculum: “so that is the thread of our curriculum structure. And the core tasks come back several times, you can develop toward master level. Core tasks 1/2/3 are addressed very prominently only one time, in their own module, so you have to show the master level immediately. But core tasks 4 and 5, and 6 in practical research, they come back three times at least ... so you can grow” (developer/teacher nr2). And: “Well, we made a rubric for the entire master ... and we said, we always assess the core tasks, in all modules. So the students develop a product and you could say, we assess the product, but we assess the three core tasks that are addressed in that product” (developer/teacher nr2). The METVE students recognize the design of the programme, but also add that even more connections could be made between the modules, to make the programme even more meaningful to them: “but you could also stress the connections between the modules ... because practical research, it would be good if you use the topic of the module about guidance, that you do research on that topic. So I would stress the connections much more, that you can use practical research in the other modules. I think, now we have three separate modules whereas there are so many connections. Now I see that, I think ... well, you

should also stress it at the start of the METVE (1st year student nr2).

The second theme illustrates how the programme design alone does not guarantee meaningfulness of assessment criteria. It needs to be purposefully designed *and* realized by the teachers. METVE students—especially the 1st year students—tell they find it difficult to look far ahead: “well, let’s be honest, we are starting students, really, we are not trying to find out what you have to do three years from now” (1st year student nr2). This teacher also realizes 1st year students cannot have an overview of the entire programme: “this is how I tell students in my module [end 2nd year], because I think at that moment they can understand. Because I think it is quite complex if you tell this at moment zero. Because they do not understand the programmatic perspective yet” (developer/teacher nr1). This theme thus also seems to show differences between the needs of 1st year and 2nd year students, which has implication for programme design and teacher activities (e.g., a full programmatic perspective might be too much to ask from 1st year students).

METVE teachers mentioned some strategies they use to foster students’ meaning making at the programme level. Teachers not only give feedback on current assignments, but also give feed forward that indicates what students have to do to improve toward the master’s level (as described in the overall rubric): “what I do, when I assess bachelor-plus level, then I write down what they have to do to reach the master’s level. So even if only bachelor-plus is required, I always add this, for the master level” (developer/teacher nr2). Also, teachers try to stimulate students to make connections between the different modules, in which they work on their core tasks and grow toward the master level: “in the entire METVE you want to guide them toward a certain level. And if you cannot see what your module contributes to this development ... if you cannot put that next to the contribution of the other modules ... well, that is not handy” (teacher nr1). And: “what we do is, we ask students at the start of a module, well bring the assessment form of the previous module. They get a lot of feedback, and we ask students what are your strong and weak points if you look at your last assignment and feedback. And what does that mean, what are you going to work on now” (developer/teacher nr3).

To work on meaning making of assessment criteria at the programme level, students need to have an impression of their own development toward the master level (theme 3). Only then, students can connect their own development to the assignments of the modules and their choices of what to work on and improve. METVE students and teachers described how they notice development and growth. For example: “You notice that they strive for quality, you see assumptions and arguments ... that they suddenly realize, why am I doing this? When I see that...” (developer/teacher nr1). And: “actually, they are too successful, that is a criterion for me ... [...] ... they get more tasks. So you notice ... at their workplace. They are taken seriously as a discussion partner, they notice more ... I am not sure whether they are more interfering with matters ... you see their workload increases” (developer/teacher nr1). The following examples show how METVE students notice their own development (or not):

“Well yes, if you have reached the required level ... or if it is only for this assignment, that is the question of course. So overall ... actually I do not have a clear impression for myself” (1st year student nr4).

“For me ... that I learn to use tacit knowledge, so I get a grip on the domain, on being a vocational teacher. Just because you notice much better what you are doing ... I notice that my work really develops, I make more deliberate choices” (2nd year student nr4).

“Well ... I maybe find the results of the assignment less valuable. What I am looking for is my performance at the workplace. I think I should act at another level in my organization. And this has nothing to do with whether I finish the assignments and get a pass. So that part, the credit points you get ... that is nice and I know I have to do it. But they don't mean too much for where I am standing now. (2nd year student nr1).

Finally, the last theme shows a number of prerequisites for meaning making of assessment criteria at the programme level. The METVE programme has ran for just two years and the teachers still need to develop a certain routine: “Now we have to develop a routine, as a team. And you have to learn to carry out the design, you have to be on the same wavelength. That is phase we are in now” (developer/teacher nr1). Developers, teachers and students agreed that they are still searching for the different possible interpretations of assessment criteria: what choices can be made, what are the different variations of good work. Teachers develop these impressions during their first (and subsequent) years of teaching, in which they encounter many variations of student work. This also raises some issues with regard to new teachers who start working as a teacher in the METVE programme, because meaning making at the programme level requires a full overview of the curriculum. In the interviews, teachers told they are better able to guide students and help them make meaning of the criteria after their first year of teaching: “well, I notice, now we do it the second time, that I can be sharper in dialogues with students ... I am better able to guide the discussions because I formed a picture of what they can choose ... you can give examples” (developer/teacher nr2). Also, to really work toward the programme goals in all modules requires that teachers have an overview of the entire curriculum, share these goals and know what is going on in other modules. In other words, as one of the developers said: “so it is a team effort and not an effort of teacher who all do their own little part” (developer/teacher nr3).

## The Contribution of Self-Regulation to Students' Meaning Making Process

For the third research question, “How can self-regulation contribute to students' meaning making process?” six themes were derived from the data, related to (1) students' own development at the module level, (2) students' own development at the programme level, (3) lack of time, (4) students' professional role after graduation, (5) making self-regulation more explicit, and (6) supporting self-regulation.

The first theme, students' own development at the module level, refers to the choices students make for their assignments. These choices are not always based on what their professional context is asking, but also on how they want to develop themselves and what they want to learn. So, at the module level, students are challenged to show ownership. As one of the teachers explained this: “We do this by asking them to find domain experts [for a specific assignment]... and what you encounter is the quality of the expert (...) that way we try ownership...” (developer/teacher nr2). Students pick up that responsibility and are sometimes proactive: “Well, I complete an assignment, hand it in and then I ask feedback. Sometimes I ask feedback from my critical friends [peers], which is often supporting, some advice together. Or feedback from the teacher and based on that feedback I get a notion of what is actually meant.” (2nd year student nr4).

The METVE teachers acknowledge how important it is that students pay attention to their development at the programme level, which is the second theme concerning self-regulation. Developer/teacher nr2 stated: “When we designed the programme, we said that in particular at Master's level, and in particular where it concerns a teachers' programme, it is very important that students are able to self-assess, and that they are able to handle such an assessment instrument [the general rubric].” Not all students are able to show self-regulation at the programme level. There are big individual differences. Teacher 2 explained: “I think that the majority [of the students], not everyone, are focused on their development at Master's level. And how they can contribute to their professional practice.” Developer/teacher nr1 also noticed that some students show quite passive behavior throughout the programme: “Sometimes we [teacher and student] joke about it, like hey, you behave like a student again, what's this? We discuss that with them, like well, you can act as a sort of passive student and the teachers says this and that,... so that is sometimes a sort of discussion, and we say that is not the way it works...” As has been stated before, the freedom students have for completing assignments the way they want can make them a bit insecure and, consequently, quite passive. Several students thus indicated that they want more explicit attention for self-regulation at the programme level. For example, 2nd year student nr2 indicated: “I think I know my strong and weak points, but we miss some sort of career counselor.” So, the programme allows for self-regulation at the programme level, but it is not explicitly addressed and supported. It is up to the individual student whether they will accept the challenge.

The third theme, lack of time, is related to what was just described. Students indicate that at the beginning of the year they were asked to write down what they wanted to accomplish, but because the programme is perceived as very time consuming and difficult to combine with a job and family, they revealed that they did not have the time to reflect on these personal learning goals. Some students are happy if they can just do what the teachers tell them and they feel that setting their own learning goals costs extra time they don't have: “Studying 20 h a week in combination with a job, that is just hard to formulate your own learning goals, etcetera. So, indeed, it is like okay, I do that module, tak

tak, preparing the face-to-face meeting for next week. It is very tight schedule, check stuff, prepare your own work, and then it is already Monday again.” (1st year student nr5).

The fourth theme, students’ professional role after graduation, was only mentioned in a few occasions. Basically, the message was that students’ professional role after graduation is hardly addressed: “[students] are really looking for their role, like how can I get this all together, and what do other people gain from me as an expert teacher at Master’s level? (...) They find that really hard, because their role is new, no-one really knows yet, so the tasks you [the expert teacher after graduation] gets, do they fit?” (developer/teacher nr2). Students and teachers recognize this: “We actually never discussed each others’ personal learning goals and never reflected on how they fit in the METVE, and how come that you have these goals, and does it have something to do with the opportunities you have at work... and then I come back to what am I going to do after I finished the METVE, for me in my higher education institute. My personal goals are really related to that” (2nd year student nr3) and “Yes, because they [students] have no idea which role they will get in the future, what they are able to do and know when they are Master’s teacher in vocational education. But we could ask them to create that image, based on what they learn now, what they think, what they get out of the programme, what kind of role they see for themselves, when they are that Master’s teacher. That will put it more in perspective, yes, maybe learn from it and a career or something, an orientation on the career (...) and getting the self-regulation from that” (teacher nr2).

The last two themes, make self-regulation more explicit and support of students’ self-regulation, are related in the sense that the teachers have an important role in this. In the METVE, there is some attention for the development of self-regulation in the design of the programme, but in reality, this is rather implicit. Teachers find it important to explicitly communicate this expectation, that is, that METVE students take up self-regulation at the programme and module level. “Yeah, I had this conversation [with a student] (...) and I heard her saying: what do I have to do? And I said, I think we offer boundaries in which learning takes place, so what do you want to learn? That is nice of METVE, students can fill that in for themselves. But I think we stated that too implicitly” (teacher nr1), and “I think that for self-regulation, that starts with expectations... what do I expect of students at the end of module XXX, what are the attainment levels, and are students allowed to show a more or lesser degree of self-regulation...” (teacher nr2), and “We have this vision, but we don’t show that explicitly, and that makes that students not always pick that [self-regulation] up, and we just say, yeah, they are Master students so...” (teacher nr2).

So, in the METVE, the development of self-regulation could have been more supported. Students are not able to develop their self-regulation at the programme level automatically. Implicitly, teachers expect students to be able to when they enroll and acknowledge that this can be improved. “Maybe we should guide that in a better way or build that up, because now we kind of let that go in my opinion, and assume that they will do that, and maybe we could say that first we take them by the hand and then... give them more freedom” (teacher nr1) and “Well, they

[students] can ask feedback on the things they are working on, and they do that, but maybe they have to ask specific feedback questions. Because in the design we said that before you ask feedback from your teacher, you have to formulate a specific feedback question. Don’t just go to the teacher and say, here is my assignment, what do think of it? (...) They find that hard. It demands a bit of self-evaluative judgment, because you have to be able to estimate what your strong and weak points are to be able to ask a good feedback question.” (developer/teacher nr3) Teachers do not always feel that they are able to support students’ self-regulation at the programme level sufficiently. On the one hand, they need time, a good overview of the curriculum, and a notion of what students will do after they are graduated. On the other hand, because the programme is still quite new, at the moment they are mainly busy with improving the quality of the content of their own module. As teacher nr3 puts it: “I always think, in the future, when I have more hold on it, next year, I can improve. Now I am less focused on... stimulating students to formulate their personal goals... I didn’t even think of that... but if I would have wanted to do that... I think let’s first just manage my own content before I go outside of that. I just can’t handle that at the moment.” (teacher nr3).

## CONCLUSION AND DISCUSSION

The research questions that guided this exploratory case study focused on fostering meaning making of assessment criteria at the module level and the programme level, and the role of self-regulation in this meaning making process. We presented a single exploratory case study in order to explore processes of meaning making of assessment criteria by curriculum designers, teachers, and students. It needs to be noted that this study was conducted in a teacher education context and the results might not apply to other higher education courses. Also, we used interviews in which the participants self-reported about their experiences with regard to the meaning making of assessment criteria, which might have affected the results.

Our study explored how meaning making takes place at the module level (research question 1) by using holistic assessment criteria which allow students to make choices within the boundaries set by the assessment criteria. Comparable to Sadler’s (1989) argument, the METVE teachers seem to value holistic assessment criteria, but the counter side may be that students seem to experience insecurity as holistic criteria provide less guidance on what is expected. In this respect, previous research on the use of rubrics (e.g., Jonsson and Svingby, 2007) shows that value of task specific rubrics. When holistic criteria are used, meaning making at the module level should thus be fostered by creating evaluative experiences, such as comparing examples, peer feedback and modeling practices by the teacher. The evaluative experiences mentioned by the METVE teachers show they go beyond telling and showing desired learning outcomes, as is recommended by several authors (cf. Fluckiger et al., 2010; Willis, 2011; Hawe and Dixon, 2014).

This study also indicates that meaning making seems to be more difficult at the programme level (research question 2).

The design of the METVE programme aims to foster meaning making at the programme level by using a general holistic rubric that is used in all modules, and by assessing the core tasks throughout the curriculum, so students can show growth (van der Vleuten et al., 2012; cf. Bok et al., 2013a). However, the design of a curriculum alone cannot ensure meaning making at the programme level. As research on evaluative judgment shows, fostering students' capacity for evaluative judgment requires pedagogic practices, in this case focused on the programme level and students' long term learning process toward the programme goals. The METVE teachers give feedback and feed forward using the general rubric and stimulate students to take feedback from one module to another. This approach to giving feedback is also advocated by Hughes et al. (2015) to stimulate student learning beyond the module level. Other teacher and student activities seem to center around creating evaluative experiences for students (Boud et al., 2018). Examples found in this study are peer assessment, dialogues and intervision activities. More research is needed, however, to explore to what extent these activities really focus students' attention on the programme goals or graduate learning outcomes. Students seem to tend to focus on the upcoming assignment and less on their development throughout the programme and beyond.

When it comes to self-regulation (research question 3), we have seen that the METVE programme (including the assessment criteria and standards) is designed in such a way that the design allows for self-regulation at the programme level. If students want to formulate and evaluate their own learning goals, they can (within the boundaries of the holistic assessment criteria). However, it is up to the individual student whether they will take on this challenge and students seem not to do this spontaneously (only guided by a teacher). Students also perceive it as something extra for which they actually don't have enough time and the freedom they have in the programme sometimes makes them feel insecure. Furthermore, self-regulation at the programme level is not explicitly supported in the METVE, for example by a study coach. That makes it more difficult for some students to create meaningfulness of assessment criteria and standards and use this meaning making process to regulate their own learning process toward the role they want to fulfill after completing the Master's programme.

In general, this study provides some practical implications for the design of higher education courses and (starting) teacher professional development in higher education when it comes to fostering meaning making of assessment criteria at the module level and the programme level. First, our case study seems to indicate some prerequisites for meaning making to happen, especially at the programme level. In order to design and carry out evaluative activities that foster meaning making at the programme level, teachers—just as students—need to develop an overview of all possible varieties of student work that fit within the holistic assessment criteria. Also, teachers need to be familiar with the entire curriculum—and not just their “own” modules—to be able to give feedback and feed forward across modules. Teaching thus becomes a team effort instead of an individual activity (cf. Jessop and

Tomas, 2017). Just as students can discuss different examples of student work to develop a more diverse picture of what quality might entail in diverse vocational situations, (starting) teachers could do the same in professional development activities and discussions when judging student work (Sadler, 1989; Boud et al., 2018).

Second, curriculum designers (in higher education) could take into account the programme perspective from the start of their design process. They could design a sequence of assessment methods that increasingly stimulate students' capacity for evaluative judgment. They could design evaluative experiences at the programme level and even beyond, by addressing the role of the students after graduation (Boud, 2000), when professionals also have to be able to judge what good work entails. Evaluative judgment concerns the evolving ability to engage with quality criteria and make informed judgments about one's own work and that of others (Boud, 2000; Sadler, 2009; Carless, 2015; Panadero and Broadbent, 2018). During their engagement in the curriculum with different assignments and activities organized by the teachers, students can gradually develop a sense of quality, like in Carless' study (2015) by presentations and peer feedback resulting in an increase in transparency by exemplifying how criteria and standards can be applied in diverse products.

Third, in order to foster meaning making of assessment criteria at the programme level, students can be stimulated to take a more active role in meaning making processes. This study revealed that students—in this case busy working students—tend to work with the assessment criteria on their own (for example by using the assessment criteria as a checklist). We believe that for students (as well as teachers), meaning making could benefit from collaborative processes like intervision, peer feedback, and dialogues (e.g., Sadler, 1989; Carless, 2015). This also implies that, despite time pressure and insecurity, higher education students should develop an attitude to deal with insecurity and work on meaning making and self-regulation in collaboration. Although it seems quite hard for students, we argue that this is necessary for students to become lifelong learners.

Finally, a discussion about analytic vs. holistic assessment criteria seems warranted. In our exploratory case study, holistic criteria were used to foster meaning making and clarity of student progress at the programme level. Holistic criteria leave room for meaning making and self-regulation (Sadler, 2009). On the other hand, assessment instruments with a more analytical perspective may be beneficial for students as well (e.g., Weigle, 2002; Jonsson and Svingby, 2007), because task specific rubrics and analytic criteria may provide more specific diagnostic information for improvement that can be used by teachers and students. Govaerts et al. (2005) found that students' experiences with regard to more analytic vs. holistic assessment criteria varied depending on their experience. Beginning students prefer analytic assessment criteria because they provide clear guidance for learning, whereas more experienced students prefer holistic criteria as they perceive analytic criteria to be checklists that do not really capture what is important in professional practice. Our study seems to indicate a similar distinction. This exploratory case study thus seems to indicate that what is needed with regard to assessment criteria might

be different at the beginning of a curriculum than at the end. Again, this seems to advocate a programme perspective on curriculum design, with more analytic assessment criteria and specific meaning making activities at the beginning, and more holistic criteria, options of task customization and peer feedback toward the end of the curriculum. Future research could take up the challenge to explore how students experience their engagement with assessment criteria throughout the curriculum. Also, because the current study was a single exploratory case study, more research is needed to further investigate the programme perspective on assessment criteria, for example on evaluative judgment and how students' capacity for evaluative judgment can be fostered throughout the curriculum, and how students' attention can be geared toward the programme goals and the future profession instead of the short-term upcoming assignment.

## REFERENCES

- Baartman, L. K. J., and Gulikers, J. T. M. (2017). "Assessment in Dutch vocational education: overview and tensions of the past 15 years," in *Enhancing Teaching and Learning in the Dutch Vocational Education System: Reforms Enacted*, eds E. de Bruijn, S. Billett, and J. Onstenk (Springer), 245–266.
- Black, P., and William, D. (1998). Assessment and classroom learning. *Assess. Educ.* 5, 7–74. doi: 10.1080/0969595980050102
- Bloxham, S., and Campbell, L. (2010). Generating dialogue in assessment feedback: exploring the use of interactive cover sheets. *Assess. Eval. Higher Educ.* 35, 291–300. doi: 10.1080/02602931003650045
- Bok, H. G., Teunissen, P. W., Favier, R. P., Rietbroek, N. J., Theyse, L. F., Brommer, H., et al. (2013a). Programmatic assessment of competency-based workplace learning: when theory meets practice. *BMC Med. Educ.* 13:123. doi: 10.1186/1472-6920-13-123
- Bok, H. G., Teunissen, P. W., Spruijt, A., Fokkema, J. P. I., van Beukelen, P., Jaarsma, D. A., et al. (2013b). Clarifying students' feedback-seeking behaviour in clinical clerkships. *Med. Educ.* 47, 282–291. doi: 10.1111/medu.12054
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Stud. Cont. Educ.* 22, 151–167. doi: 10.1080/713695728
- Boud, D., Ajjawi, R., Dawson, P., and Tai, J. (2018). *Developing Evaluative Judgement in Higher Education. Assessment for Knowing and Producing Quality Work*. London: Routledge. doi: 10.4324/9781315109251
- Boud, D., and Falchikov, N. (2006). Aligning assessment with long-term learning. *Assess. Eval. Higher Educ.* 31, 399–413. doi: 10.1080/02602930600679050
- Brooks, J., McCluskey, S., Turley, E., and King, N. (2015). The utility of template analysis in qualitative psychology research. *Qual. Res. Psychol.* 12, 202–222. doi: 10.1080/14780887.2014.955224
- Brown, E., and Glover, C. (2006). "Evaluating written feedback," in *Innovative Assessment in Higher Education*, eds C. Bryan and K. Clegg (London: Routledge), 81–91.
- Carless, D. (2015). Exploring learning-oriented assessment processes. *High. Educ.* 69, 963–976. doi: 10.1007/s10734-014-9816-z
- Clark, I. (2012). Formative assessment: assessment is for self-regulated learning. *Educ. Psychol. Rev.* 24, 205–249. doi: 10.1007/s10648-011-9191-6
- Conway, R. (2011). *Owning Their Learning: Using Assessment for Learning to Help Students Assume Responsibility for Planning, (some) Teaching and Evaluation*. Teaching History, 51.
- Dannefer, E. F., and Henson, L. C. (2007). The portfolio approach to competency-based assessment at the Cleveland clinic lerner college of medicine. *Acad. Med.* 82, 493–502. doi: 10.1097/ACM.0b013e31803ead30
- Fluckiger, J., Vigil, Y. T. Y., Pasco, R., and Danielson, K. (2010). Formative feedback: Involving students as partners in assessment to enhance learning. *Coll. Teach.* 58, 136–140. doi: 10.1080/87567555.2010.484031
- Govaerts, M. J. B., Van der Vleuten, C. P. M., Schuwirth, L. W. T., and Muijtjens, A. M. M. (2005). The use of observational diaries in in-training evaluation: student perceptions. *Adv. Health Sci. Educ.* 10, 171–188. doi: 10.1007/s10459-005-0398-5
- Gulikers, J., and Baartman, L. K. J. (2017). *Doelgericht Professionaliseren: Formatieve Toetspraktijken Met Effect! Wat DOET de Docent in de Klas? [Goal Directed Professionalisation: Formative Assessment Practices With Effect! What Do Teachers Do in the Classroom]*. Netherlands Initiative for Education Research. Report number: NRO PPO 405-15-722.
- Hartley, P., and Whitfield, R. (2011). The case for programme-focused assessment. *Educ. Dev.* 12, 8–12.
- Hawe, E. M., and Dixon, H. R. (2014). Building students' evaluative and productive expertise in the writing classroom. *Assess. Writ.* 19, 66–79. doi: 10.1016/j.asw.2013.11.004
- Heeneman, S., Oudkerk Pool, A., Schuwirth, L. W., Vleuten, C. P., and Driessen, E. W. (2015). The impact of programmatic assessment on student learning: theory versus practice. *Med. Educ.* 49, 487–498. doi: 10.1111/medu.12645
- Hughes, G., Smith, H., and Creese, B. (2015). Not seeing the wood for the trees: developing a feedback analysis tool to explore feed forward in modularised programmes. *Assess. Eval. High. Educ.* 40, 1079–1094. doi: 10.1080/02602938.2014.969193
- Jessop, T., and Tomas, C. (2017). The implications of programme assessment patterns for student learning. *Assess. Eval. High. Educ.* 42, 990–999. doi: 10.1080/02602938.2016.1217501
- Jonsson, A., and Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educ. Res. Rev.* 2, 130–144. doi: 10.1016/j.edurev.2007.05.002
- Kearney, S. (2013). Improving engagement: the use of 'authentic self-and peer-assessment for learning' to enhance the student learning experience. *Assess. Eval. High. Educ.* 38, 875–891. doi: 10.1080/02602938.2012.751963
- Lokhoff, J., Wegewijs, B., Durkin, K., Wagenaar, R., Gonzales, J., Isaacs, A. K., et al. (2010). *A Tuning Guide to Formulating Degree Programme Profiles. Including Programmes Competences and Programme Learning Outcomes*. Bilbao; Groningen; Hague: Nuffic/TUNING Association.
- Lorente, E., and Kirk, D. (2013). Alternative democratic assessment in PETE: an action-research study exploring risks, challenges and solutions. *Sport Educ. Soc.* 18, 77–96. doi: 10.1080/13573322.2012.713860
- Nicol, D. J., and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud. High. Educ.* 31, 199–218. doi: 10.1080/03075070600572090
- Panadero, E., and Broadbent, J. (2018). "Developing evaluative judgment: self-regulated learning perspective," in *Developing Evaluative Judgement in Higher Education. Assessment for Knowing and Producing Quality Work*, eds D. Boud, R. Ajjawi, P. Dawson, and J. Tai (London: Routledge), 81–89.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Faculty Ethics Review Board (FERB) of the Faculty of Social and Behavioural Sciences of Utrecht University. Since this research project is a simple observational study that does not involve any interventions but just comprises interviews, the study has not been subject to review by an ethical committee. This is also in accordance with the recommendations of the FERB. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

- Panadero, E., and Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assess. Educ.* 21, 133–148. doi: 10.1080/0969594X.2013.877872
- Parr, J. M., and Limbrick, L. (2010). Contextualising practice: hallmarks of effective teachers of writing. *Teach. Teach. Educ.* 26, 583–590. doi: 10.1016/j.tate.2009.09.004
- PASS position paper (2012). *The Case for Programme Focused Assessment*. Available online at: [www.pass.brad.ac.uk](http://www.pass.brad.ac.uk)
- Price, M., Carroll, J., O'Donovan, B., and Rust, C. (2011). If I was going there I wouldn't start from here: a critical commentary on current assessment practice. *Assess. Eval. High. Educ.* 36, 479–492. doi: 10.1080/02602930903512883
- Rust, C., Burkill, S., and Whitfield, R. (2012). *Programme Assessment Strategies. International Perspectives. Examples from outside the UK*. Available online at: [www.pass.brad.ac.uk](http://www.pass.brad.ac.uk)
- Rust, C., Price, M., and O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assess. Eval. High. Educ.* 28, 147–164. doi: 10.1080/02602930301671
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instruct. Sci.* 18, 119–144. doi: 10.1007/BF00117714
- Sadler, R. (2009). "Transforming holistic assessment and grading into a vehicle for complex learning," in *Assessment, Learning and Judgement in Higher Education*, ed G. Joughin (Dordrecht: Springer), 45–63.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assess. Educ.* 14, 281–294. doi: 10.1080/09695940701591867
- van der Vleuten, C. P., Schuwirth, L. W. T., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K. J., et al. (2012). A model for programmatic assessment fit for purpose. *Med. Teach.* 34, 205–214. doi: 10.3109/0142159X.2012.652239
- Weigle, S. C. (2002). *Assessing Writing*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511732997
- Wiliam, D. (2011). What is assessment for learning? *Stud. Educ. Eval.* 37, 3–14. doi: 10.1016/j.stueduc.2011.03.001
- Willis, J. (2011). Affiliation, autonomy and assessment for learning. *Assess. Educ.* 18, 399–415. doi: 10.1080/0969594X.2011.604305
- Yin, R. K. (2014). *Case Study Research Design and Methods*. Los Angeles: Sage.
- Zimmerman, B. J. (2000). "Attainment of self-regulation: a social cognitive perspective," in *Handbook of self-regulation*, eds M. Boekaerts, P. R. Pintrich, and M. Zeidner (San Diego, CA: Academic Press), 13–39. doi: 10.1016/B978-012109890-2/50031-7
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: an overview. *Theory Pract.* 41, 64–70. doi: 10.1207/s15430421tip4102\_2

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Baartman and Prins. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.