This is a pre-print of the article that was published as

Authenticity is in the Eye of the Beholder: Student and Teacher Perceptions of

Assessment Authenticity

Judith T. M. Gulikers [a*], Theo J. Bastiaens [b, c], Paul A. Kirschner [c], and Liesbeth Kester [c]

Open University of the Netherlands

[a] Education and competence studies, Wageningen University, Wageningen, The Netherlands

[b] Institute for Educational Science and Media Research, Fernuniversität in Hagen, Hagen, Germany.

[c] Open University of the Netherlands, Heerlen, The Netherlands

---

[*] Corresponding Author. Email: judith.gulikers@wur.nl

Abstract

In VET in the Netherlands, learning and working are integrated from the start. Authentic

assessments are used during competence-based VET curricula to achieve correspondence

between learning and working. The premise behind this study is that authenticity is subjective

and that perceptions of assessment authenticity influence student learning for the assessments.

It examines if students and teachers differ in their perceptions of the authenticity of various

assessment characteristics. Subsequently it investigates if freshman and senior students, who

differ in their amount of practical experience, differ in their perceptions of assessment

authenticity. The main findings were that teachers rated most assessment characteristics as

more authentic than students did, while freshman and senior students did not differ in their

perception of authenticity. Implications deal with communicating about and developing

authentic assessment in the eyes of both students and teachers to stimulate students'

professional skills development during a VET curriculum

Keywords: Authentic assessment; Curriculum; Student perception; Teacher perception;

Teacher professionalization; Vocational education and training

Authenticity is in the eye of the beholder: student and teacher perceptions of assessment authenticity

The issue

Authenticity, defined as *resembling students' (future) professional practice* (Gulikers, Bastiaens, & Kirschner, 2004), is a crucial characteristic of competence-based assessments in Vocational Education and Training (VET) colleges (Boud, 1995; Gulikers, Bastiaens, & Kirschner, 2006a; Segers, Dochy, & Cascallar, 2003; Wesselink, Biemans, Mulder, & van der Elsen, 2007a). The premise of this study is that authenticity is, at lest partly, subjective; the perception of authenticity might change as a result of for example amount of professional experience or age (Gulikers et al., 2004; Radinsky, Boullion, Lento & Gomez, 2001; Roelofs & Terwel, 1999; Stein, Isaacs, & Andrews, 2004). We argue that for authentic assessments to reach their potential in positively stimulating student learning and better preparing them for the labour market, it is imperative that *students* perceive their assessments as authentic. This study examines the differences between student and teacher perceptions of assessment authenticity by examining if two assessments that are developed by teachers to be authentic, are indeed perceived as such by freshman and senior students.

Authentic assessments in Dutch competence-based VET

To better prepare students for the labour market, the Dutch government obliges VET colleges to have competence-based curricula and assessments by 2010. These competence-based (VET) programmes integrate learning and working in professional practice from the start (Wesselink et al., 2007a). This means that they strive for a correspondence between what students have to do during learning or assessment and what students are expected to do during internships or after finishing their school (Boud, 1995, Messick, 1994; Stein et al., 2004; Tillema, Kessels, & Meijer, 2000). Authentic assessments are thought to help bring learning, assessment and working closer together. Following this reasoning, this study defines

authenticity in terms of its *resemblance to students' (future) professional practice* (Gulikers et al., 2004). By creating this resemblance, authentic assessment are thought to show students the link between learning and working in practice, thereby directing their learning towards developing professionally relevant skills (Herrington & Herrington, 1998; Lizzio & Wilson, 2004a; McDowell, 1995; Wesselink et al., 2007a).

A problem however is that authenticity is not purely objective, but dependent on a person's perception of the authenticity of the assessment (Gulikers et al. 2004; 2006a). In addition, these student perceptions of assessment characteristics seem to mediate the influence of an assessment on student learning (Boud, 1995; Gijbels, 2005; Scouller, 1997). An important question thus is if students (users of an assessment) and teachers (developers of the assessment) perceive assessment authenticity in the same way. Differences or similarities with respect to assessment authenticity between students and teachers or students of different years of study might have important implications for using authentic assessments during a curriculum.

Subjective authenticity and the role of perceptions

The educational goal of authentic assessments is to stimulate deep learning activities and the development of more professionally relevant skills or competences (Boud, 1995; Gulikers et al., 2006a; Tillema et al., 2000). Unfortunately, this relationship between assessment and learning is not that straightforward. Learning is influenced by assessment in three ways (Boud, 1995, p. 36), namely (a) by the intrinsic or objective qualities of the assessment; (b) by a teacher's interpretation of the to be assessed material. A teacher translates the material to be assessed into a certain format and select assessment tasks appropriate for the subject and the specific learning goals; and (c) by a student's interpretation of the task at hand and the context of the assessment. Previous research (Entwistle, 1991; Gijbels, 2005; Scouller & Prosser, 1994; Scouller 1997; Struyven, Dochy, & Janssen, 2003;

Van Rossum & Schenk, 1984) has shown that especially the third element, being *students'*

perceptions of the assessment characteristics, is crucial for determining learning. For

example, when students perceive the assessment as measuring recall of factual information,

students employ a surface study strategy that seems suitable for learning factual information

by heart (Gijbels, 2005; Scouller 1995; Scouller & Prosser, 1994; Struyven 2005)

With respect to assessment authenticity, student perceptions seem to be influential as

well (Gulikers, Bastiaens, Kirschner, Kester, 2006b). The aforementioned definition that

assessment authenticity depends on the degree of resemblance between the assessment and

students' professional practice might make is seem as if authenticity is an objective construct,

but authenticity is not an 'objective' quality as such. Something is only authentic with respect

to something else, for example a situation, place or profession (Gulikers et al., 2006a;

Honebein et al. 1993; Radinsky et al., 2001). Whether a person sees an assessment as being

authentic depends on the reference point that person has in mind against which the

authenticity is measured. In addition, a person's perception of authenticity can also change.

This can be the result of the amount and kind of practical experience, schooling or age

(Honebein et al; Lizzio & Wilson, 2004a; Petraglia, 1998). This means that what one person

perceives as being authentic is not necessarily authentic in the eyes of someone else (Gulikers

et al., 2004; 2006a). In reality, thus, "authenticity is in the eye of the beholder". If it is true

that students' perceptions of assessment authenticity drives their learning then this indicates

that before an authentic assessment can positively influence learning, it is imperative that the

learner *perceives* the assessment as being authentic (Radinsky et al., 2001).

Differences in perceptions of authenticity

Several studies have shown that there are differences between teacher and student

perceptions of a learning environment or assessment (Boud, 1995; Ngar-Fun, 2005;

MacLellan, 2001; Sambell & McDowell, 1998; Wesselink, Biemans & Mulder, 2007b).

Teachers often use an assessment to send a 'message' to students about what kind of learning is required, but students' perception of this message is not always in line with the teacher's intention (Boud, 1995; Sambell & McDowell). In other words, students create their own 'hidden curriculum'; they interpret the learning environment and assessment practices in their own way, which in turn drives their learning. Lizzio and Wilson (2004a), for example, showed that *students' perceptions* of relevance of the to-be-developed or to-be-assessed skill for their future work drives their willingness and interest in acquiring that particular skill.

Previous studies suggest that it cannot be automatically assumed that what teachers see as being authentic, and thus what they make use of in the lessons or assessments that they develop, is also perceived as being authentic by students (e.g. Cummings & Maxwell, 1999; Roelofs & Terwel, 1999). Moreover, the authors previously found that VET-teachers view authenticity in a much more detailed way than VET-students (Gulikers et al., 2006a), which might have important implications for authentic educational practices. In light of the effect of authenticity on learning, moderated by *student* perceptions of this authenticity (Gulikers et al., 2006b), it is important that assessments are developed that students perceive as being authentic. To this purpose, this study examines if two assessment that are developed by teachers to be authentic, are indeed perceived as such by their students.

Besides the possible differences between students and teachers, students with different amounts of experience in professional practice might also differ in how they perceive assessment authenticity. Because students in VET in the Netherlands start doing internships from the very beginning of their studies, they are gaining much professional practice experience during their studies. Senior students, thus, have had a lot of practical experience, while freshman do not. Lizzio and Wilson (2004b) argued that students with little professional experience have unrealistic expectations about work and work roles, while seniors, having more professional experience, might have changed perceptions of work and work lives

(possibly in the direction of more realistic expectations). These differences might influence what both student groups perceive as authentic assessments. This, in turn, might have important consequences for designing and using authentic assessments during a VET curriculum.

This study

The main focus of this study is on examining differences in assessment authenticity perceptions of teachers versus students and freshman versus senior students. A validated and renewed five-dimensional framework (5DF) for assessment authenticity (Gulikers et al., 2006a) is used as a tool for describing the 'objective' authenticity of the assessments used in this study and for examining assessment authenticity from the student and teacher perspectives. This framework is grounded in theory on authenticity and authentic assessment (Gulikers et al., 2004), validated by students and teachers (Gulikers et al., 2006a) and adapted to these findings. Moreover, this framework has recently been adopted by the Dutch Association for VET Colleges (MBO-raad, 2007) as a useful instrument for describing, improving and quality assuring the authenticity of assessments (MBO-raad, 2007). This framework argues that five assessment characteristics influence the degree of authenticity of the assessment as a whole. The five assessment characteristics can be described as follows:

1. Task. The assessment assignment that defines the content of the assessment

2. Physical context. The environment in which students have to perform the assessment task

3. Social context. The interaction (im)possibilities during the assessment

4. Form. The assessment method, independent of the content

5. Criteria. The characteristics of the performance (product/process) that are value

The rationale behind this framework is that these five characteristics can resemble professional practice to a more or lesser extend (Gulikers et al., 2004). Thus, an assessment can be made more or less authentic on five continuous scales. This framework makes it

possible to describe and examine the resemblance between professional practice and these five assessment characteristics.

The research questions of this study are: (1) how do students and teachers evaluate the authenticity of an assessment that is developed to be authentic by the teachers and do they agree?, and (2) do freshman students, with little professional practice experience, and senior students, with a lot more practical experience, differ in how authentic they perceive the assessment characteristics to be?

Method

Participants

A group of final students ($n = 118$; Mean age = 19.16; SD = 1.14) and a group of first year students ($n = 66$; Mean age = 18.13; SD = 1.67) studying Social Work at a VET college enrolled in this study. Seniors would graduate within four months, while freshman students started their studies six month earlier. Students were studying Social Work in a competency-based learning environment combined with authentic assessments. Both groups differed in their amount and kind of practical experience. Freshman students had been working in one professional setting for one day a week. Senior students had completed various internships, which differed from one day a week to ten weeks full time. In addition, 17 teachers of the freshmen program and 19 teachers of the senior program participated in this study. They were involved in the authentic assessment as a developer and an assessor or role-player. There was no overlap between teacher groups.

As a precondition for examining student perceptions, the researchers explicitly selected student groups that were familiar with the kind of assessment used in the study for two reasons. First, Struyven (2005) shows that if students are unfamiliar with an assessment method, their preference for this assessment is lower than for assessment methods they are familiar with. But after having experienced the new assessment method once, their preference

increased significantly. It is possible that the same kind of process holds for students' perceptions of the assessment, which makes the perceptions after the second experience with the assessment a more reliable one. Second, if students are confronted with something new and unfamiliar in their learning environment, they first have to adapt to this change (Gibbs, 1992). Evaluating the learning environment might be affected by the students' ability or willingness to adapt to the changes. Evaluating an element of the learning environment that students are already used to increases the likelihood of evaluating the element of interest. Both student groups were familiar with the kind of assessment used in this study. For the freshman students it was the second time that they performed this kind of assessment, for the senior students it was the seventh time.

Materials

This study made use of two existing assessments at a VET institute for social work, to assure the ecological validity of the study. Obviously, the used assessments could not be completely identical as one was for freshmen students and one for seniors, however they were both developed along the same assessment format and they were designed to be authentic assessment for the target group. This assured that the assessments were as comparable as possible. The only difference between the assessments was the topic.

The topic of the freshman students' assessment was 'dealing with conflict situations'. This was one of the main competences that students had to acquire during the course 'orientation towards your own possibilities'. It consisted of a case describing a situation where a student's pupil was not allowed to take part in the institute-festivities because the family's religious background. The child was angry with his parents because he wanted to join the other children in the festivities and have fun with them while the parents and their religion forbid this. During the assessment, the students had to solve this problem during a role-play in which the teacher played the mother of the child.

The topic of the senior students' assessment was 'applying for a job' on one of three vacancies for just graduated social workers. This was one of the main competences that students had to acquire during the course 'the social work organisations: working with policy'. During the assessment they had to take part in a job interview with the teacher playing the role of the employer.

The 'objective' authenticity of the assessments was described according to the five dimensions of the 5DF and is shown in Table 1. Two researchers independently scored the authenticity of the five dimensions based on a document analysis of the assessment material and one of the reviewers observed several student performances during the assessments. The rating of the authenticity was based on the degree of resemblance between the elements of the five dimensions and professional practice. This resulted in scores on a 5-point scale ranging from very high degree of resemblance ('++') to almost no resemblance ('- -').

*** INSTERT TBALE 1 HERE ***

Objectively speaking, both assessments had the same authenticity for both groups, except for the social context. The individual, one-on-one assessment context is authentic in the case of the 'applying for a job' assessment, since a job interview is mostly done individually in real life as well. In the case of 'dealing with conflicting situations', an individual assessment is less authentic, since in real life, a social worker might choose to deal with this problem together with a colleague and when children are involved they are likely to meet with both parents.

A competency-based instructional period of nine weeks preceded both authentic assessments. During eight weeks, students worked in groups on critical professional problem situations. They had to set learning goals focusing on knowledge as well as skills/attitudes. During this training phase, students had to carry out several formative assessments. These were all role-play assignments, based on a social work related problem situation, that student

had to carry out with other students. The summative assessment was based on a selection of course objectives that was translated into the assessment criteria. The assessment criteria for the summative assessment were conveyed to students one week prior to the assessment in which students were freed from obligatory educational activities.

The questionnaire for measuring perceptions of authenticity of various assessment characteristics was based on the five dimensions of the 5DF and adapted to previous findings with testing this questionnaire with students and teachers in VET (Gulikers et al., 2006a). The scales examined the authenticity of the five dimensions (task, physical context, social context, form, criteria). The 24-items of the questionnaire all assessed the perception of the resemblance of one of the assessment characteristics with (future) professional practice ('The task of this assessment prepared me for my future professional life'). The items were scored on a 5-point Likert scale ranging from 'totally disagree' to 'totally agree', resulting in a score for the perceived degree of resemblance between the assessment characteristics and professional practice. All scales, except for the social context scale, had a reasonable internal consistency, shown in Cronbach's alpha ranging from .70 to .82. The social context scale was left out of further analysis.

Procedure

During one week, all students took part in the assessment. Students filled in all questionnaires directly after finishing the assessment. The teacher questionnaire was almost identical to the student perception questionnaire, except that the word 'I' was replaced with the words 'the student'.

Analysis

To examine if the groups agreed on the authenticity of the five assessment characteristics, two measurements were used. Fist, one-sample t-tests were used in which the means of the groups were individually compared to the mean score of the rating scale (value

3) to find out if the four groups rated the authenticity of the assessment characteristics above or below average. Second, ANOVA tests were used to examine if and how the groups differed in their ratings of the authenticity of the characteristics. Games-Howell post hoc tests were used, since this is an appropriate test in the case of different group sizes.

Results

Table 2 shows the mean scores on the authenticity scales for the four groups.

*** INSERT TABLE 2 HERE ***

In line with the objective scoring of the authenticity of the assessment characteristics (see Table 1), all groups perceived the *physical context* as the least authentic characteristic of the assessments (Table 2). In addition, the one-sample t-tests showed that the four groups valued the physical context as less then averagely authentic (lower than three), with the student groups rating the physical context significantly lower than three ($t(65) = 2.88$, $p < 0.01$ for first year students; $t(117) = 5.53$, $p < 0.01$ for final year students). The *task*, *form* and *criteria* of the assessment were valued as more than averagely authentic in all groups ($p < 0.01$) except for the final year students group in which the rating of the task did not deviate from the median score of 3 ($t(117) = 1,47$, *ns*.). In the objective rating of the authenticity of the assessment (table 1) the task, form and criteria were scored as (highly) authentic as well.

ANOVA tests showed significant differences between the four groups on the *task*, *form* and *criteria* dimension ($F(3, 216) = 12.86$, $p < 0.01$; $F(3, 216) = 12.73$, $p < 0.01$; $F(3, 216) = 17.28$, $p < .01$ respectively). Games-Howell post hoc tests showed a similar picture for the three dimensions (for Means and Standard Deviations see table 2): (a) in all cases, teachers scored significantly higher than students ($p < 0.01$) (b) both student groups did not differ from one another; and (c) the two teacher groups did not differ from each other. The four groups did not differ in their perception of the authenticity of the *physical context* ($F(3, 216) = 1.38$, *ns*.).

Conclusion and discussion

At a general level, students and teachers agreed about the resemblance of the assessment characteristics and professional practice being above or below average. Their ratings at this general level were also in line with the objective ratings based on the five-dimensional framework (Table 1). The five-dimensional framework seems to differentiate appropriately between various characteristics that are important in developing authentic assessments in VET. A closer look at student and teacher perceptions however showed that there are important differences that should be taken into account. Teachers perceived most characteristics of assessment as more related to professional practice (i.e. as more authentic) than students did. On the other hand, freshman and senior students did not differ in their perceptions of authenticity of the same kind of assessment.

The findings give food for thought for further research, but first some considerations need to be taken into account when interpreting the results of this study. First, a five-dimensional framework (Gulikers et al., 2004) was used to evaluate the authenticity of the assessments. This was used to score the 'objective' authenticity of the assessment, but this can never be completely objective, since it will always be an appraisal done by a person. By using two independent raters and by doing a document analysis as well as an actual observation of the assessment, the rating was thought to be as objective as possible. Moreover, the adoption of the five-dimensional framework by the Dutch Association of VET Colleges (MBO-raad, 2007) suggests the practical relevance and face validity of the instrument for VET education and its possibility to evaluate the authenticity of an assessment. Second, one can question the degree to which a role-play assessment (i.e., the authentic assessments in this study) can be called authentic with respect to professional practice. However, the overall ratings of all students and teachers did not suggest doubting the overall authenticity of the assessment used in this study.

Differences and similarities in perceptions of teachers versus students

Two reasons can explain the differences between student and teacher perceptions and especially the finding that teachers perceived most characteristics as significantly more authentic than students. First, as a result of the gap between teachers' beliefs and their actual assessment practices (Orrell, 2003; Verhoeven & Verloop, 2002; Wiggins, 1989), the actual practices might be much less authentic than teachers think they are, while students only see the actual practices. There is no one-on-one relationship between what teachers believe that they are implementing and what they actually implement. Second, as a result of more and different kinds of experience in professional practice, teachers are likely to have a different idea of what professional practice looks like than students do (Lizzio & Wilson, 2004a; Radinsky et al., 2001; Roelofs & Terwel, 1999), while teachers develop assessments according to their ideas of authenticity without taking students' ideas into account.

With respect to the first explanation, several studies showed that there often is a gap between what teachers think they are doing and what they are actually doing (Orrell, 2003; Maclellan, 2001; Ngar-Fun, 2005). Wiggins (1989) and Maclellan even argued that the greatest disjunction between beliefs about good assessment practice and actual practice was associated with authentic assessment. Teachers often think that their assessment practices are more authentic than they actually are (Verhoeven & Verloop, 2002). This issue is further complicated by the change that VET colleges in the Netherlands, and to various extends also in other parts of the world (Mulder & Wiegel, & Collins, 2007), are going through towards competence-based education and assessment. Teachers have to change their current, often traditional, beliefs about testing towards the ideas of competence-based assessments which strongly focuses on using authentic assessments that resemble professional practice (Wesselink et al., 2007a). This is a development that does not go automatically and requires time (Biemans, Nieuwenhuizen, Poell, Mulder & Wesselink, 2004). In this light, Cummings

and Maxwell (1999) argued that many educational practices are implemented in the name of authentic assessment while in the end these often are old, existing assessment that are "dressed up" with some "realistic" elements. This is regularly seen in educational practice in the Netherlands at this moment. Teachers are obliged to develop authentic assessments, but they have not yet internalised or accepted the underlying beliefs about these kind of assessments. Cooper (1994), however, showed that resulting "dressed up" assessments are often not perceived as authentic by students and as a result hampered their learning.

With respect to the second explanation for different student and teacher perceptions, it is widely accepted that previous experiences colour the way current learning environments are perceived (Biggs, 1989; Birenbaum, 2003; Sternberg, 1999; Wesselink et al., 2007b). Previous study (Gulikers et al., 2006a) already suggested that teachers, having more educational and practical experience than students, have a much more detailed frame of reference for interpreting authentic assessment than student do. Teachers develop an assessment according to what *they* think is authentic for the professional field (Huang, 2002; Petraglia, 1998). Petraglia called this "pre-authentication" (p. 53) which is "the attempt to make learning materials and environments correspond to the real world prior to the learner's interaction with them". This then reflects the real world as teachers see it, while from a learners point of view, this might not be authentic at all, since they might have a different perception of what the real world involves (Lizzio & Wilson, 2004a; Roelofs & Terwel, 1999; Stein et al., 2004). The pitfall is that while teachers are often the ones to develop the authentic assessment, *students'* perceptions of this assessment determine their learning. The discrepancy between student and teacher perceptions might signal that students do not have an accurate or complete picture of what professional practice entails, but is can also be the result of outdated or unrealistic beliefs of teachers, as many teachers haven't been working in jobs they are educating for.

Both explanations would support Cummings and Maxwells (1999) or Petraglias (1998) invitation to problematise the concept 'authentic', and especially, what it means in different contexts and for different people (e.g., different stages of VET curricula). Petraglia argued that a problem with authenticity is that it is not communicated explicitly. It is such an intuitive concept that people do not feel the need to be explicit about it since 'everybody knows what we are talking about'. The results of this study suggest that it is not so obvious that everybody sees authenticity in the same way. Both discussed explanations suggest that explicating tacit beliefs of both teachers and students about authentic assessment and professional practice might be an important step for developing better authentic assessments that stimulate students' learning and motivation at various stages of their education. For example, Lizzio and Wilson (2004a) argued that in order to motivate students' learning of professionally relevant skills, teachers should pay particular attention to explicitly discussing the relevance of the authentic assessment practices for students' future working lives, since students' interest in developing skills is largely dependent on their perceived relevance of the skill to their future work. This supports an other crucial characteristic of new assessments, next to authenticity, namely a higher involvement of students in developing assessments (e.g, Birenbaum et al., 2006).

Freshman versus senior student perceptions

Contrary to the expectation, freshman and senior students did not differ in their perception of assessment authenticity. Previous studies also expected to find differences between students of different years of study, but these expectations were not confirmed either (Handal & Hofgaard Lycke, 2005; Winning et al., 2005). Handal and Hofgaard Lycke compared freshman to senior students with respect to their way of learning and to the kind of competences that they thought were important for their future work. This study showed that with respect to both these characteristics, students did not differ. On the other hand, when the

senior students were tested again after one year of working, their ways of learning and the competences that they thought were important in professional live had changed. This might mean that students' perceptions and way of learning are relatively stable during their years of studying and more internship experience does not influence this, while after finishing school, work experience seems to drastically change their ideas. Pena (1997) also argued that when students graduate and enter the professional field, they often experience a 'reality shock', because they experience that the real world of work is still (completely) different from what they expected while studying. These findings suggests that students remain in a 'school-based frame of reference' for interpreting their assessments that is changed after leaving school and entering the labour market. This suggests, just as Boud in 1990 already argued, that there is still a big gap between learning / assessment and working, which can explain why first and final year students do not differ, while students and teachers do differ.

To conclude, explicating how teachers and various student groups perceive authenticity or authentic assessment seems to be a crucial step in developing authentic assessments that appropriately resemble professional practice in the eyes of both students and teachers. It needs to be recognised that both students as well as teachers can have misconceptions. Students might have a too narrow, or too idealistic picture of professional practice, while teachers can have outdated beliefs (Radinsky et al. 2001). Moreover, changing towards competence-based assessment, which all VET colleges in the Netherlands are obliged to, is a struggle in itself and a quest for both students and teachers (Biemans et al., 2004; Wesselink et al., 2007b). This requires both professionalization of teachers as well as helping students understand and appreciate the ideas of competence-based education with authentic assessments. Communication between students and teachers about what authenticity and professional practice means at different stages of a VET curriculum should facilitate (collaborative) development of authentic assessments and thereby decrease the gap between

learning and working (Boud, 1995). If, as suggested by the results of this study, students'

perceptions are indeed stable during their years of study and turn out to be difficult to change

during studying, it might be more valuable to collaboratively develop one effective kind of

authentic assessment to use during a VET curriculum than putting a lot of effort, time and

energy in developing all different kinds of authentic assessments during a VET curriculum.

Acknowledgements

Notes on contributors

Judith T. M. Gulikers, Chairgroup of Education and competence studies of the Wageningen University and Research Centre of the Netherlands

Paul A. Kirschner, and Liesbeth Kester, Educational Technology Expertise Center of the Open University of the Netherlands.

Theo J. Bastiaens is now at the Ruud de Moor Centre of the Open University of the Netherlands and at the Institute for Educational Science and Media Research of the Fernuniversität in Hagen.

References

Biemans, H., L. Nieuwenhuis, R. Poell, M. Mulder, and R. Wesselink. 2004. Competence-based VET in the Netherlands: Background and pitfalls. *Journal of Vocational Education and Training* 56:523-538.

Biggs, J. 1996. Enhancing teaching through constructive alignment. *Higher Education* 32; 47-364.

Birenbaum, M. (1996) Assessment 2000: Towards a pluralistic approach to assessment. In *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge*, ed. M. Birenbaum and F J.R.C. Dochy, 3-29. Boston: Kluwer Academic Publishers.

Birenbaum, M., K Breuer, E. Cascallar, F. Dochy, Y. Dori, J. Ridgeway, and R. Wiesemes. 2006. A learning integated assessment system. *Educational Research Review* 1:61-69.

Black, P. and D. Wiliam. 1998 Assessment and classroom learning. *Assessment in Education* 5: 7-74.

Boud, D. 1990. Assessment and the promotion of academic values. *Studies in Higher Education* 15: 101-111.

Boud, D. 1995. Assessment and learning: contradictory or complementary? In *Assessment for Learning in Higher Education*, ed. P. Knight, 35-48. Londen: Kogan Page.

Cooper, B. 1994 Authentic testing in mathematics? The boundary between everyday and mathematical knowledge in national curriculum testing in English schools. *Assessment in Education: Principles, Policy and Practice* 1: 143-166.

Cummings, J. J. and G.S. Maxwell. 1999. Contextualising authentic assessment. *Assessment in Education: Principles, Policy and Practice* 6: 177-194.

Dochy, F. (2001) A new assessment era: different needs, new challenges. *Research Dialogue in Learning and Instruction* 10: 11-20.

Dochy, F.J.R.C. and L. McDowell. 1998 Assessment as a tool for learning. *Studies in Educational Evaluation* 23: 279-298.

Entwistle, N. J. 1991. Approaches to learning and perceptions of the learning environment. Introduction to the special issue. *Higher Education* 22: 201-204.

Gijbels, D. 2005 Effects of new learning environments. Taking students' perceptions, approaches to learning and assessment into account. PhD diss., University of Maastricht, The Netherlands.

Gulikers, J.T.M., Th.J. Bastiaens, and P.A. Kirschner. 2004. A five-dimensional framework for authentic assessment. *Educational Technology Research and Development* 52: 67-85.

Gulikers, J.T.M. , Th. J. Bastiaens, and P.A. Kirschner. 2006a. Authentic assessment, student and teacher perceptions: the practical value of the five-dimensional framework. *Journal of Vocational Education and Training* 58:337-357.

Gulikers, J.T.M., Th.. J. Bastiaens, P.A. Kirschner, and L. Kester. 2006b. Relations between student perceptions of assessment authenticity, study approach and learning outcome. *Studies in Educational Evaluation* 32:381-400.

Handal, G. and K. Hofgaard Lycke. 2005. *From higher education to professional practice: learning among students and novice professionals*. Paper presented at the European Association for Research on Learning and Instruction Conference, August 23-27, Nicosia, Cyprus.

Herrington, J. and A. Herrington. 1998. Authentic assessment and multimedia: how university students respond to a model of authentic assessment. *Higher Educational Research & Development* 17: 305-322.

Honebein, P.C., T.M. Duffy, and B.J. Fishman. 1993. Constructivism and the design of learning environments: Context and authentic activities for learning. In T. M. Duffy, J.

Lowyck and D. H. Jonassen, eds, *Desinging Environments for Constructive Learning*, pp 88-108. Berlin: Springer-Verslag.

Huang, H. M. 2002. Towards constructivism for adult learners in online learning environments. *British Journal of Educational Technology* 33: 27-37.

Lizzio, A. and Wilson, K. 2004a. First-year students' perceptions of capability. *Studies in Higher Education* 29: 109-128.

Lizzio, A. and K. Wilson. 2004b. Action learning in higher education: an investigation of its potential to develop professional capability. *Studies in Higher Education* 29: 469-488.

Maclellan, E. 2001. Assessment for learning: the differing perceptions of tutors and students. *Assessment and Evaluation in Higher Education* 26: 307-318.

MBO-raad. 2007. Competentiegericht beoordelen in het MBO. Handreikingen voor het middenmanagement [Competence-based assessment in VET. Guidelines for middle management]. De Bilt, the Netherlands: MBO-raad.

McDowell, L. 1995 The impact of innovative assessment on student learning. *Innovations in Education and Training International* 32: 302-313.

Messick, S. 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher* 23: 13-23.

Morgan, D. L. 1997. *Focus groups as qualitative research.* London: Sage Publications.

Mulder, M., T. Weigel, and K. Collins. 2007. The concept of competence concept in the development of vocational education and training in selected EU member states. A critical analysis. *Journal of Vocational Education and Training* 59: 67-88.

Ngar-Fun, L. *Hong Kong Academics' and Students' Perceptions of Assessment Purposes and Practices*. The Hong Kong Institute of Education. http://www.ied.edu.hk/loap/index.html.

Pena, E. 1997. Great expectations: the reality of the workplace. *Australian Journal of Career Development* 6: 25-32.

Petraglia, J. 1998. *Reality by Design: The Rhetoric and Technology of Authenticity in Education*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Radinsky, J., L. Boullion, E.M. Lento, and L.M. Gomez. 2001. Mutual benefit partnership: a curricular design for authenticity. *Journal of Curriculum Studies* 33: 405-430.

Roelofs, E. and J. Terwel. 1999. Constructivism and authentic pedagogy: state of the art and recent developments in the Dutch national curriculum in secondary education. *Journal of Curriculum Studies* 31: 201-227.

Sambell, K. and L. McDowell. 1998. The construction of the hidden curriculum: Messages and meanings in the assessment of student learning. *Assessment and Evaluation in Higher Education* 23: 391-402.

Scouller, K. M. and M. Prosser. 1994. Students' experiences in studying for multiple choice question examinations. *Higher Education* 19: 267-279.

Scouller, K. M. 1997. Students' perceptions of three assessment methods: Assignment essay, multiple choice question examination, short answer examination. *Research and Development in Higher Education* 20: 646-653.

Segers, M., F. Dochy, and E. Cascallar. 2003. *Optimising New Modes of Assessment: In Search of Qualities and Standards*. Dordrecht: Kluwer Academic Press.

Stein, S. J., G. Isaacs, and T. Andrews. 2004. Incorporating authentic learning experiences within a university course. *Studies in Higher Education* 29: 239-258.

Sternberg, R. J. 1999. *Cognitive psychology (2nd ed.)*. Fort Worth, TX: Harcourt Brace College Publishers.

Struyven, K. 2005. The Effects of Student-activated Teaching/Learning Environments of Students' Perceptions, Student Performance and Pre-service Teachers' Teaching. PhD diss., University of Leuven, Leuven, Belgium.

Tillema, H. H., J.W.M Kessels, and F. Meijers. 2002. Competences as building blocks for integrating assessment with instruction in vocational education: a case study from The Netherlands. *Journal of Curriculum Studies* 25: 265-278.

Van Rossum, E. J. and S.M. Schenk. 1984. The relationship between learning conceptions, study strategies and learning outcome. *British Journal of Educational Psychology* 54: 73-83.

Verhoeven, P. and N. Verloop. 2002. Identifying changes in teaching practice: Innovative curricular objectives in classical language and the taught curriculum. *Journal of Curriculum Studies* 34: 91-102.

Wesselink, R., H. J. A. Biemans, M. Mulder, and R. Van der Elsen. 2007a. Competence-based VET as seen by Dutch researchers. *European Journal of Vocational Training* 40:38-51.

Wesselink, R., H. J. A. Biemans, and R.H. Mulder. 2007b. Perceptions of CBL-environments in VET. Paper presented at the European Conference of Educational Research, September 19-21, in Ghent, Belgium.

Wiggins, G. 1989. Teaching to the (authentic) test. *Educational Leadership* 46: 41-47.

Winning, T., L. Elaine, and G. Townsend. 2005. Student experiences of assessment in two problem-based dental curricula: Adelaine and Dublin. *Assessment and Evaluation in Higher Education* 30: 489-505.

Table 1.

The objective authenticity based on the five dimensions of assessment authenticity (Authors,

2004)

Assessment dimension	Description	Value
Task	Case description of a situation that is representative of students' current internship or near professional future. Based on a core competence of Social Work as described in collaboration with the work field	++
Physical context	In school, in a classroom. An unknown teacher is the role-player. The timeframe of the assessment is ten minutes, which is not realistic since in real life the talks would stop when finished. no resources are available, which are also not likely to be used in the real life situation	-
Social context	One-on-one. First year: This might not resemble professional practice, since in real life at least both parents will be present and the student might involve another colleague.	-
	Final year: an application is always done individually	+
Form	Role-play. Doing these kinds of talks are an important part of being a social worker making this an authentic demonstration of competence	++
Criteria	Criteria are developed in collaboration with professional practice, scored on a three-point scale, made known one week before the assessment and most are directed towards observable behaviour or talk. However, two criteria mainly focus on knowledge that students have to express	+

Table 2.

Mean scores on the perception scales of the four groups

	Senior students (n = 118)	Freshman students (n = 66)	Senior teachers (n = 17)	Freshman teachers (n = 19)
Task	3.10 (.77)	3.21 (.48)	3.92 (.67)	3.86 (.59)
Physical context	2.53 (.92)	2.76 (.68)	2.85 (1.00)	2.63 (.87)
Form	3.31 (.74)	3.41 (.59)	4.22 (.36)	3.93 (.50)
Criteria	3.20 (.62)	3.26 (.42)	4.05 (.53)	3.86 (.63)