**Deliverable 2.4: Literature review on workplace-based**

<table>
<thead>
<tr>
<th>Delivery month Annex I:</th>
<th>M10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual delivery month</td>
<td>M10</td>
</tr>
<tr>
<td>Lead participant: UMCU</td>
<td>Work package: 2</td>
</tr>
<tr>
<td>Version: 1.0</td>
<td>Nature: R</td>
</tr>
<tr>
<td></td>
<td>Dissemination level: PU</td>
</tr>
</tbody>
</table>

**Project coordinator**

**Dr. Marieke van der Schaaf**

Utrecht University  
Faculty of Social and Behavioral Sciences  
Department of Education  
PO Box 80.140  
3508TC Utrecht  
The Netherlands

Telephone: +31 (0)30 253 4944  
Email: M.F.vanderSchaaf@uu.nl
1. Executive Summary ................................................................. 3
2. Introduction ............................................................................. 3
   2.1 Background ........................................................................ 3
   2.2 Scope of the deliverable ..................................................... 3
3. Content .................................................................................... 3
4. Conclusion ............................................................................... 22
5. References ............................................................................. 26
6. Tables and Figures ................................................................. 26
7. History of the document ......................................................... 26
   7.1 Document history ............................................................... 26
   7.2 Internal review history ....................................................... 26
1. Executive Summary
This deliverable aims to extract from the pertinent literature in the domains of medical education (part 1), veterinary education (part 2) and teacher training (part 3) the most valuable approached and instruments to inform entrustment decisions and feedback, suitable to be represented in learning analytics based visualization connected to an electronic portfolio. The large numbers of instruments may be roughly grouped into the following categories.

<table>
<thead>
<tr>
<th>Practice unrelated instruments</th>
<th>Practice related instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td></td>
</tr>
<tr>
<td>• Written or electronic knowledge testing</td>
<td>• Short practice observation</td>
</tr>
<tr>
<td>• Case-based discussion</td>
<td></td>
</tr>
<tr>
<td>Skill</td>
<td></td>
</tr>
<tr>
<td>• Simulation testing</td>
<td>• Short practice observation</td>
</tr>
<tr>
<td>Attitude</td>
<td></td>
</tr>
<tr>
<td>• Simulation testing</td>
<td>• Long practice observation</td>
</tr>
<tr>
<td>Products</td>
<td></td>
</tr>
<tr>
<td>• Product evaluation</td>
<td>• Product evaluation</td>
</tr>
</tbody>
</table>

The finding of the systematic reviews that are summarized in this deliverable will be combined with the empirical findings to be reported in deliverables 2.3 and 2.4.

2. Introduction

2.1 Background
With reference to Deliverable 2.1 this deliverable is intended to provide an overview from the literature of sources.

2.2 Scope of the deliverable
This deliverable is relevant for workplace assessment in the domains and institutions that contributed for the WATCHME project, but also more broadly for these domains at other universities and countries, and possibly for other professional domains. As the foundation is laid for technology-based feedback and entrustment decision-making as an approach to assessment, it will stimulate learning analytics development that is yet in its infancy in these domains of higher education.

3. Content

Part 1:
Markers and instruments in the medical workplace to inform in trainee progression, feedback and entrustment decisions, as reported in the literature.
Authors: Nienke Wisman-Zwarter (lead), Reinier Hoff, Marieke van der Schaaf, Johannes Kaldeway, Ylva Holzhausen, Olle ten Cate, Wilton van de Klei
Introduction

Assessment in medical education is vital. The last decades, medical education changed from a master-apprentice structure to a learner centered, competency-based approach. This change initiated the need for more appropriate assessment tools. More recently, Entrustable Professional Activities (EPAS) are being developed for a number of undergraduate and postgraduate medical educations. EPAs make it possible to follow the progression of a trainee or resident in their professional development. Different assessment tools are used for each EPA. The aim of this review is to gain an overview of the current available assessment tools used in performance assessment.

Method

We searched ERIC and Medline for articles written in English and published between January 1990 and December 2014. See table 1 for our search query.

Inclusion criteria
We included original research on performance assessment and entrustment decision making concerning both undergraduate and/or postgraduate medical education.

Exclusion criteria
Review articles, expert opinions, and consensus surveys were excluded. We also excluded articles on non-medical (e.g. pharmacy or dental training) or non-academic (e.g. nursing, physiotherapy, and midwifery) subjects.

Review
The literature search yielded 418 articles after the initial practical screen of which 133 articles were relevant for this review. We decided to disregard articles containing Virtual Reality or Augmented Reality Simulation as research topic due to their very specific field of application. Articles are categorized according to their assessment setting, namely Educational, Clinical, or Simulation. See table 2 for the complete overview of all references.

Table 1. Search Query

<table>
<thead>
<tr>
<th>AND</th>
<th>&quot;formative&quot; AND (&quot;assessment&quot; OR &quot;assessments&quot; OR &quot;evaluation&quot; OR &quot;evaluations&quot; OR &quot;feedback&quot;) OR &quot;summative&quot; AND (&quot;assessment&quot; OR &quot;assessments&quot; OR &quot;evaluation&quot; OR &quot;evaluations&quot; OR &quot;feedback&quot;) OR &quot;feedback&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;performance” AND (&quot;assessment&quot; OR &quot;assessments&quot; OR &quot;evaluation&quot; OR &quot;evaluations&quot; OR &quot;feedback&quot;) OR &quot;performance based” AND (&quot;assessment&quot; OR &quot;assessments&quot; OR &quot;evaluation&quot; OR &quot;evaluations&quot; OR &quot;feedback&quot;) OR &quot;authentic” AND (&quot;assessment&quot; OR &quot;assessments&quot; OR &quot;feedback&quot;)</td>
</tr>
</tbody>
</table>

4
Results

A total of 133 articles were found eligible for this review. Of these articles 50% was published in 2010 or later. Previously, assessment tools were designed for usage in an educational setting (skill stations or classroom setting). However, since the year 2000 assessment tools specifically adapted for the clinical setting are emerging and ultimately cover for half of all assessment tools found in this review. See figure 1 for further information.

Simulation setting
Publications on assessment tools designed for a simulation setting are arising since 2005. Especially team skills and non-technical skills or crisis resource management (CRM) skills are subject for assessment tools in this setting.

**Educational setting**
The Objective Structured Clinical Examination (OSCE) is in 1990 the first described assessment tool as well as the most modified of all assessment tools in the three settings. OSCEs are adapted to be used in different specialties (e.g. undergraduate medical education clerkships and postgraduate pediatrics training) and used for both formative and summative assessment.

**Clinical setting**
The multisource assessment tool is more and more used in the clinical setting to obtain information on all competencies of a trainee. This information complements the information gained from the more practical assessment methods like the mini Clinical Examination Exercise (mini-CEX) or the Objective Structured Assessment of Technical Skills (OSATS). New developments are made in assessment tools for assessment of less tangible competencies and skills, e.g. chairing case conferences (case Conference Assessment Tool (cCAT)).

<table>
<thead>
<tr>
<th>What is assessed</th>
<th>Tool</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation setting</td>
<td>Surgical skill, Non-technical skills (CRM), Communication, Teamwork</td>
<td>Simulation</td>
</tr>
<tr>
<td>Clinical performance</td>
<td>Simulated Clinical Examination (SCE)</td>
<td>11</td>
</tr>
<tr>
<td>Educational setting</td>
<td>Medical knowledge, Clinical Judgment, Reflection, Decision making</td>
<td>Case-based discussion (CBD)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Formative Assessment Case Studies (FACS) (Structured) Oral examination (SOE)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Professional competence</td>
<td>Peer assessment protocol (PAP)</td>
</tr>
<tr>
<td></td>
<td>Medical knowledge, Communication, Clinical Judgment, Confidence, Decision making</td>
<td>Objective structured clinical examination (OSCE)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Remote examiner OSCE (reOSCE)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Objective, structured communication assessment of residents (OSCAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Teaching performance</td>
<td>Assessment tool for resident performance as clinical teachers</td>
</tr>
<tr>
<td>Clinical setting</td>
<td>Procedural performance</td>
<td>Operative performance rating system (OPRS)</td>
</tr>
<tr>
<td>Deliverable D2.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td><strong>Surgical skill</strong></td>
<td>Procedure-based assessment (PBA)</td>
<td>57 58 59 60</td>
</tr>
<tr>
<td></td>
<td>Direct Observation of Procedural Skills (DOPS)</td>
<td>61 62</td>
</tr>
<tr>
<td></td>
<td>Global Procedural Skills Evaluation (GPSE)</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Task-specific checklist</td>
<td>64 65 66 67 68</td>
</tr>
<tr>
<td></td>
<td>Intra-operative video recording assessment</td>
<td>69 70</td>
</tr>
<tr>
<td></td>
<td>Structured Assessment of Microsurgery Skills (SAMS)</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Endoscopic Sinus Surgery Competence Assessment Tool (ESSCAT)</td>
<td>72</td>
</tr>
<tr>
<td><strong>Technical skill</strong></td>
<td>Crowd Sourced Assessment of Technical Skills (CSATS)</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Objective Structured Assessment of Technical Skills (OSATS)</td>
<td>74 75 76 77 78 79 80</td>
</tr>
<tr>
<td><strong>Technical skills, Nontechnical skills</strong></td>
<td>Operative performance assessment tool</td>
<td>81 82 83</td>
</tr>
<tr>
<td></td>
<td>Surgical Training and Assessment Tool (STAT)</td>
<td>84 85</td>
</tr>
<tr>
<td><strong>Debriefing</strong></td>
<td>Objective structured assessment of debriefing (OSAD)</td>
<td>86</td>
</tr>
<tr>
<td><strong>Chairing case conference</strong></td>
<td>case Conference Assessment Tool (cCAT)</td>
<td>87</td>
</tr>
<tr>
<td><strong>Written communication,</strong></td>
<td>Sheffield Assessment Instrument for Letters (SAIL)</td>
<td>88</td>
</tr>
<tr>
<td><strong>Communication,</strong></td>
<td>Patient Note Scoring Rubric</td>
<td>89 90</td>
</tr>
<tr>
<td><strong>Professionalism</strong></td>
<td>StudentPEP patients' evaluation</td>
<td>91 92</td>
</tr>
<tr>
<td></td>
<td>Communication Assessment Tool (CAT)</td>
<td>93</td>
</tr>
<tr>
<td><strong>Clinical performance,</strong></td>
<td>Interaction Cards (IC)</td>
<td>94 95</td>
</tr>
<tr>
<td><strong>CanMEDs roles‡</strong></td>
<td>Daily evaluation forms</td>
<td>96</td>
</tr>
<tr>
<td><strong>Clinical performance,</strong></td>
<td>Clinical Performance Evaluation System</td>
<td>97 98</td>
</tr>
<tr>
<td><strong>ACGME core competencies‡</strong></td>
<td>Clinical-performance biopsy (CPB) system</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Multisource feedback</td>
<td>100 101 102 103 104 105 106 107 108 109 110 111 112</td>
</tr>
<tr>
<td><strong>Clinical performance,</strong></td>
<td>Systematically observed clinical encounter (SOCE)</td>
<td>113</td>
</tr>
<tr>
<td><strong>Clinical performance,</strong></td>
<td>Clinical observed performance evaluation</td>
<td>114 115</td>
</tr>
<tr>
<td><strong>Clinical performance,</strong></td>
<td>Preanesthesia Consultation (PAC) Scoring</td>
<td>116</td>
</tr>
</tbody>
</table>
Workplace-based e-Assessment Technology for Competency-based Higher Multi-professional Education

Deliverable D2.4

<table>
<thead>
<tr>
<th>Checklist</th>
<th>Team work, Clinical performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Team Performance assessment</td>
</tr>
<tr>
<td></td>
<td>Teamwork Mini-Clinical Evaluation Exercise (T-MEX)</td>
</tr>
<tr>
<td></td>
<td>Mini-Clinical Evaluation Exercise (Mini-CEX)</td>
</tr>
<tr>
<td></td>
<td>Professionalism Mini-Evaluation Exercise (P-MEX)</td>
</tr>
<tr>
<td></td>
<td>R-I-M-E (Reporter-Interpreter-Manager-Educator) evaluation</td>
</tr>
<tr>
<td></td>
<td>Online Peer-evaluation of Handoffs</td>
</tr>
<tr>
<td></td>
<td>Structured Clinical Observation (SCO)</td>
</tr>
<tr>
<td></td>
<td>Direct observation of clinical skills (DOCS)</td>
</tr>
<tr>
<td></td>
<td>Unannounced standardized patients (SPs)</td>
</tr>
<tr>
<td></td>
<td>117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133</td>
</tr>
</tbody>
</table>

± ACGME core competencies: Patient Care, Medical Knowledge, Practice-based Learning and Improvement, Interpersonal and Communication Skills, Professionalism, and Systems-based Practice
‡ CanMEDs roles: Medical Expert, Communicator, Collaborator, Manager, Health Advocate, Scholar, and Professional

**Discussion and conclusion**

The number of assessment tools found in the simulation setting was influenced by our decision to disregard articles containing Virtual Reality or Augmented Reality Simulation as research topic. There were an additional 26 articles, published since 2005, on very specific simulation tools, including comparative studies on two or more different simulators. In recent years there has been an effort to develop and validate tools for the assessment of skills and competencies in both undergraduate and postgraduate medical education. The modification of existing assessment tools is a commonly used procedure. This results in a substantial amount of quite similar assessment tools with different names and abbreviations. A higher degree of conformity would allow a more clear overview of all available assessment tools in medical education.

**References**


Part 2:
How assessment tools provide insight in students’ performance in the clinical workplace: A systematic review of the veterinary literature

Authors: Chantal Duijn, Harold Bok, Erik Driessen, Mira Mandoki & Wim Kremer

Introduction

Workplace learning is seen as an essential component of clinical education by the veterinary profession (Magnier, 2011). To become a competent veterinary professional, students must be given opportunities to train and to develop in a supportive environment with a focus on results (Bok, 2014). Becoming a clinician within any health care profession involves not only learning clinical knowledge and skills, but also developing appropriate attitudes and behaviors (Mossop, 2013). Current veterinary curricula tend to identify competency domains or roles as areas specific to the profession in which students need to develop their performance (Bok 2011, NAVMEC). Juxtaposing the competency domains to the core activities of the profession could provide a blueprint for feedback and assessment that ensures that all core professional activities are identified while embracing all relevant competency domains for the profession (Scheele 2008).

In the clinical workplace, feedback is usually provided directly after observing students that have performed authentic tasks (Cantillon, 2008). This provides students with information that they can use to consolidate or to improve their performance. Ideally, performance-relevant feedback also offers insight into the steps they can take to achieve improvements (Shute, 2008). Providing continuity in this process of feedback and reflection fosters students’ competency development. In current veterinary clinical education a variety of different assessment tools for both formative and summative purposes are used to provide performance-relevant feedback.

Within educational literature an overview of veterinary assessment tools, used to provide students with feedback on their performance in the clinical workplace, is lacking. Therefore the aim of this study is to conduct a systematic review of the different assessment tools used in veterinary clinical education.

Research question

Which assessment tools, described in veterinary education literature, provides insight in students’ performance in the veterinary clinical workplace?

Objectives

1. Which assessment tools provide insight in students’ performance?
2. How do these tools provide insight in students’ performance?

For instance by using a numeric scale, self-reflection or narrative feedback.
**Methods**

**Study design**

The systematic review was based on the BEME guidelines. This provided the researchers with clear instructions for a systematic, logical and explicit appraisal of available information to determine the best evidence relating to this issue. 


**Search strategy**

A comprehensive search strategy was developed, to identify, summarize and synthesize all existing evidence regarding the assessment tools. The aim of this systematic review was to get an overview of the method, workability and effectiveness of each assessment tool.

The databases that were used to identify relevant studies were chosen based on previous research by Grindley et al.

- CAB Abstracts
- Scopus
- Web of Science
- Medline (via PubMed)

Included articles for the literature review: all articles till the end of the review period

Search modes: find all search terms and MeSH terms

Apply related words: yes

Also search within the full text of the articles: yes

Limit results: full text and peer reviewed

**Search terms**

The search terms were divided into three different subjects and synonyms

1. Veterinary education and synonyms: veterinary educations/ veterinary teaching/ veterinary study/ veterinary studies/ veterinary school/ veterinary schools/ veterinary curriculum/ veterinary curricula
2. Clinical clerkship and synonyms: clinical clerkships/ workplace/ workplaces/ workplace-based/ workplace based/ clinical setting/ clinical settings/ clinical setup/ clinical setups clinical site/ clinical sites
3. Student and synonyms: students /undergraduate/ graduate/ graduates/ intern/ interns/ resident/ residents/ trainee/ trainees

The search strategy applied the same text words in all databases. However, it used varied but equivalent subject headings because they differ between databases.

**Screening and selection of studies**

The titles and abstracts generated from the electronic database searches were collated in the RefWorks reference management database. The first author (CD) inspected all the titles and
abstracts and, when in doubt, read full articles and exclude those that obviously did not meet the inclusion criteria. In case of doubt, a second author (HB) was involved. The inclusion criteria were listed in figure 1. All articles were published in the English language, so the research team was able to read all articles without the need for a translation (Hammick, 2010).

**Figure 1 Inclusion criteria**

**First review**

Included articles were read fully by the first author on the same inclusion criteria as described in figure 1. The focus was on the appropriateness of study design and analysis, and how well the study was executed.

The reviewer had rated in discussion with the second author:

- The appropriateness of the study design to answer the research question and objectives.
- How well the assessment tool was implemented.
- The appropriateness of the data analysis

The first review of the literature was conducted and consensus was reached (with a second author) about which articles had to be included. The relevant assessment tools from each study are described in the table below.

**Results**
An overview of assessment tools used in the veterinary medical education

<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
<th>Journal</th>
<th>Year</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adams, C.I.; Kurtz, S.</td>
<td>Coaching and feedback-enhancing communication teaching and learning in veterinary practice settings.</td>
<td>Journal of Veterinary Medical Education</td>
<td>2012</td>
<td>A framework for organizing and leading feedback sessions in communication programs. This approach is based on the premise that feedback should not be a mini-lecture in which an expert or a peer essentially evaluates the learner’s performance. With AOBAs, instead of giving feedback coaches engage learners in a more interactive, participatory exchange of opinions and ideas.</td>
</tr>
<tr>
<td>2</td>
<td>Bok, H.G.; Tenissen, P.W.; Favier, J.R.P.; Wierbroek, N.J.; Thyeve, J.F.; Brouwer, J.; Haarhuis, J.C.; Van Beek, P.; Van der Vleuten, C.P.; Jaarsma, D.A.</td>
<td>Programmatic assessment of competency-based workplace learning: When theory meets practice</td>
<td>BMC Medical Education</td>
<td>2013</td>
<td>Used to assess performance of practicing health professionals to improve the quality of practice. Is an intensive, observed assessment on the clinical workplace. The tool is used to assess aspects of the clinical encounter such as history-taking, physical examination, clinical decision-making, professionalism, and overall clinical competence. The mCEX is a useful formative tool for promoting observation and feedback, a longitudinal evaluation.</td>
</tr>
<tr>
<td>3</td>
<td>Borden, L. J. N.; Adams, C.L.; Ladner, L.D.</td>
<td>The use of standardized clients in research in the veterinary clinical setting.</td>
<td>Journal of Veterinary Medical Education</td>
<td>2008</td>
<td>The scenario method is used to assess the student's ability to identify and respond to client concerns.</td>
</tr>
<tr>
<td>4</td>
<td>Dale, V. H. M.; Pierce, S. E.; May, S.A.</td>
<td>Benefits and limitations of an employer-led, structured logbook to promote self-directed learning in the clinical workplace.</td>
<td>Journal of Veterinary Medical Education</td>
<td>2013</td>
<td>The use of a self-directed learning tool to assess the student's ability to use technology effectively.</td>
</tr>
<tr>
<td>5</td>
<td>Dawson, Susan D.; Miller, Tess; Goddard, Sally F.; Miller, Lisa Maag</td>
<td>Impact of Outcome-Based Assessment on Student Learning and Faculty Instructional Skills</td>
<td>Journal of Veterinary Medical Education</td>
<td>2013</td>
<td>A tool to assess students' clinical competencies in the clinical setting by self-assessment and feedback from a supervisor. There are nine clinical competencies essential to provide a taxonomy of skill development.</td>
</tr>
<tr>
<td>6</td>
<td>Hafen, M.; Drake, A.A.S.; Rush, R.; Nelson, S.C.</td>
<td>Using Authentic Client Interactions in Communication Skills Training: Predictors of Proficiency</td>
<td>Journal of Veterinary Medical Education</td>
<td>2013</td>
<td>Video feedback consultations were made from a typical student-clinician interaction includes obtaining a patient health history, conducting a physical examination, and discharging the patient. To evaluate specific communication skills, research developed questions were used. Ten questions assess communication skills such as nonverbal skills and ability to deliver questions, organize discharge, and maintain professionalism.</td>
</tr>
<tr>
<td>7</td>
<td>Hecker, K. G.; Norris, J.; Cee, J.B.</td>
<td>Workplace-based assessment in a primary-care setting.</td>
<td>Journal of Veterinary Medical Education</td>
<td>2011</td>
<td>A brief, structured assessment by a professional of a student working in a case history and physical exam within the clinical setting. The primary purpose is to provide structured feedback to the learner. An assessment evaluates a student’s clinical performance (e.g., history-taking, physical examination, communication skills, professionalism) using a standardized feedback during a five-clinical encounter (typically 15 minutes). The assessor then provides the student feedback based on the competencies meant to be assessed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Direct Observation of Procedural Skills (DOPS)</td>
<td></td>
<td></td>
<td>An observation much like the mini-CEX, with a main emphasis on procedural (technical) skills and communication skills. This assessment is done by a trained individual who could be a veterinarian or a technician with training experience. As with the mini-CEX, students should be assessed during several procedures by multiple assessors and there is the expectation of immediate feedback to the student. This assessment can be done with most procedural skills so there is no defined time requirement for the procedure.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSF</td>
<td></td>
<td></td>
<td>Used to assess performance of practicing health professionals to improve the quality of practice. Within a primary health care setting, MSF has potential to provide meaningful information because it consists of collecting multiple observations by many people across several domains over the course of a rotation. MSF consists of a self-evaluation form and assessment instruments completed by some or all of the following groups: colleagues or students, supervisors, technical staff, and clients.</td>
</tr>
</tbody>
</table>
8. Karpsha, M.; Lloyd, J.W.; Peterson, F.; Derksen, F.J.  
Practice-based education at Michigan State University  
Journal of Veterinary Medical Education  2005  
Student progress report  
To get insight in students' performance during the practice-based ambulatory program a student progress report filled in by the practitioner and student at the end of each week. This checklist is an aid to both veterinarian and student to help focus the remainder of the learning experience. At approximately mid-rotation, campus liaisons are encouraged to call the practitioner for an update on the student's progress. Using this system has helped address potential problems early with appropriate corrective actions.

Workplace-based assessment instruments in the health sciences  
Journal of Veterinary Medical Education  2012  
mCEx  
The tool is used to assess aspects of the clinical encounter such as history taking, physical examination, clinical decision making, professionalism, and overall clinical competence. The mCEx is a useful formative tool for promoting observation and feedback, a longitudinal evaluation.

10. May, S.A.; Head, S.D.  
Assessment of technical skills: best practices  
Journal of Veterinary Medical Education  2010  
Objective Structured Practical Examination (OSPE)  
To create an extended series of standard stations that could provide a more comprehensive overview of student skills and, at the same time, reduce patient variability between students. By creating a checklist of maneuvers that each student completed either successfully or not, such as feeling the pulse and counting the rate on a watch, obtains a reduced examiner variability.

11. Root, K.; Kustritz, M.V.; Molgaard, J.K.; Rendahl, A.  
Comparison of student self-assessment with faculty assessment of clinical competence  
Journal of Veterinary Medical Education  2011  
Self-assessment protocol  
Self-assessment during clinical rotations is used as a measure of the levels of competence that are required to meet accreditation standards. A certain specific measures to ensure the accuracy of student self-assessment must be employed, including active faculty mentoring of student self-assessment, student goal-setting and reflection, and availability of subsequent opportunities for practice.

Web-based documentation of clinical skills to assess the competency of veterinary students  
Journal of Veterinary Medical Education  2011  
Objective Structured Practical Examination (OSPE)  
Is used to assess and confirm student competency to perform a specific set of basic technical skills in a controlled setting with teaching animals or models. Objective is station-to-station practical examinations that evaluate students' competency to perform basic technical skills in a controlled setting. The stationary assessor evaluates a large number of students performing a single skill, resulting in consistent and reliable evaluations.

mCEx  
Is an intensive, observed patient encounter. Students conduct a history and physical examination and receive immediate feedback in seven competencies. The system is an effective educational tool.

DOPS  
Is defined as the assessment of a trainee performing a procedural skill on a patient in a clinical setting. The DOPS differs from the mini-CEx in that the DOPS assesses the competency to perform a technical procedure rather than the ability to obtain a history and perform a physical examination. In general, the time needed to assess a technical skill in a DOPS program includes the time to perform the skill as well as an additional third of this time to prepare and deliver feedback. The DOPS method has been judged by supervisors and trainees as a practical and fair assessment tool with substantial contributions to professional growth and positive interactions between instructors and students.
References


Dawson, S.D.; Miller, T.; Goddard, S.F.; Miller, L.Maag Impact of Outcome-Based Assessment on Student Learning and Faculty Instructional Practices Journal of Veterinary Medical Education 2013, 40 (2): 128-38.


Part 3

Defining performance-based indicators for assessing teacher competence in the WatchMe project

Authors: Edgar Krull, Äli Leijen, Bert Slof, Marieke van der Schaaf
Note: The text below is a copy of a section that is embedded in a more general literature review, included in Deliverable 2.2.1

This performance-based approach for assessing student teachers’ teaching skills is advocated by the international team of WATCHME (2014). It is developed on the basis of Dutch teacher competence requirements (SBL, 2004) by defining related indicator performances for teacher education needs. To this end the original list of seven competences was shorten to five professional roles (actually extended by one additional role) and those in their turn were specified through descriptions of tasks or activities (also, in terms of entrustable professional activities – EPA’s) that the student teacher is expected and entrusted to perform without supervision. The whole assessment tool consists of 8 frames (tables) from which the essence of five that are relevant for defining performance indicators is exemplified shortly in the Table 1.

A more detailed analysis of indicator performances by of performance levels reveals that not only issues of the construct validity of the competence requirements as a model of good teaching in regard of indicators performances arise in the Dutch school context (for what this

<table>
<thead>
<tr>
<th>Frame</th>
<th>Title</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Matrix mapping tool</td>
<td>1) Lists 5 types of professional roles, provides descriptions of professional activities related to these roles, and lists 11 tasks to be performed as evidences of being able to implement these roles. The professional roles are: (1) designer, supervisor, and evaluator of learning activities; (2) manager of the work environment; (3) Pedagogue; (4) member of professional community; and (5) Manager of own professional development.</td>
</tr>
<tr>
<td>2</td>
<td>Curriculum mapping tool</td>
<td>Defines requirements for study phases Curriculum_internship 1 and Curriculum_internship 2. From student completing phase one implementation of roles of the pedagogue and member of the professional community is not expected.</td>
</tr>
<tr>
<td>3</td>
<td>Performance indicators</td>
<td>States for tasks to perform indicator activities (2–4 for each task) that evidence coping with tasks related to implementing professional roles</td>
</tr>
<tr>
<td>4</td>
<td>Performance levels</td>
<td>Defines four performance levels for the tasks evidencing meeting of the roles’ requirements [Level 1 (starting); Level 2 (sufficient); Level 3 (good); Level 4 (excellent)]</td>
</tr>
</tbody>
</table>

1 An EPA is ‘a critical part of professional work that can be identified as a unit to be entrusted to a trainee once sufficient competence has been reached’. Competence is thus translated and made manageable in terms of the tasks or activities that can be safely entrusted to someone who has shown the required ability (Mulder, Ten Cate, Daalder, et al., 2010; Ten Cate, 2013). One of many innovations of the WATCHME project is introduction of minimal quality criteria for professional activities. The identification and definition of these criteria for teaching are subject of another study.
model was initially developed) but also the question of appropriateness of these performance indicators for assessing teaching skill in school context of another country. For example, in a country like Estonia national curricula for general education have a much more prescriptive nature than in the Netherlands (see e.g. Education, Audiovisual and Culture Executive Agency P9 Eurydice, 2012). Consequently, Dutch student teachers are expected to be ready for more independent professional work than their Estonian colleagues. This becomes even more evident when analyzing the indicator performances expected from Dutch student teachers by completing their teaching practice at schools (see Table 2).

This means that relatively specific criteria for performance-based assessment of teaching developed in the conditions of one country for being effective in the context of another country with different education traditions should be validated for this context. One possible way for adapting the scoring rubric to specific school context is to use Delphi method for deciding whether a specific performance indicator should be accepted, revised or removed. The construct validation of content standards for teaching students research skills used by Van der Schaaf and Stokking (2011) can serve as a prototype approach here.

### Table 2. Defining performance indicators by 4 performance levels for task 1 (Formulates a vision of the subject content and the subject didactics) in the WATCHME draft teacher evaluation matrix.

<table>
<thead>
<tr>
<th>Role/task</th>
<th>Performance indicators</th>
<th>Performance Level 1</th>
<th>Performance Level 2</th>
<th>Performance Level 3</th>
<th>Performance Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role: Designer, supervisor and evaluator of learning activities (subject didactician and supervisor of the learning process)</td>
<td>1. The teacher does/does not formulate (self-formulated) learning goals in connection with specific content (subject content/didactic competences).</td>
<td>The teacher takes over the learning goals of others and the course book and occasionally stops to think about the cohesion between the set of learning goals and the specific subject content.</td>
<td>The teacher formulates his/her own learning goals which partially match those of the specific subject content.</td>
<td>The teacher formulates his/her own learning goals which partially meet SMART.</td>
<td></td>
</tr>
<tr>
<td>Task 1: Sets learning goals for the whole curriculum and specific</td>
<td>2. The teacher does/does not make use of SMART (specific, measurable, acceptable, realistic and time related) formulated</td>
<td>The teacher does not check if the set learning goals are SMART formulated.</td>
<td>The teacher regularly checks if the set learning goals are SMART formulated.</td>
<td>The teacher formulates his/her own learning goals which partially meet SMART.</td>
<td></td>
</tr>
<tr>
<td>lessons.</td>
<td>learning goals (subject content/didactic competences).</td>
<td>guidelines.</td>
<td>lines.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------------</td>
<td>--------------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. The teacher does/does not take into consideration the starting situation of students when formulating learning goals (subject content/didactic competences).</td>
<td>The teacher incidentally stops to think about the consistency between the set of learning goals and the starting situation of the students</td>
<td>The teacher regularly checks if the learning goals match the starting situation of the students</td>
<td>The teacher formulates his/her learning goals which partially match with the measured starting situation of students</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Also, the issue of reliability of documenting (measurement) cannot be ignored, as the identification of performance levels would heavily depend on the context and personal interpretation of definitions. For example, the performance criterion “the teacher incidentally stops to think about the consistency between the set of learning goals and the starting situation of the students” for identifying the performance level might be subject of very different interpretations. The fact that student teachers’ performance-based assessment is typically based on information collected from different sources (lesson plans, lesson observations and recordings, conferences etc.) additionally emphasizes a need for clear and unambiguous definition of performance level indicators.

4. Conclusion
The literature reviews have yielded a large number of instruments to and sources for markers in the health care and classroom workplaces to inform feedback and entrustment decisions for learners. For health care contexts these may be summarized preliminarily in the following groupings:

<table>
<thead>
<tr>
<th>Practice unrelated instruments</th>
<th>Practice related instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td></td>
</tr>
<tr>
<td>• Written or electronic knowledge testing</td>
<td>• Short practice observation</td>
</tr>
<tr>
<td>• Case-based discussion</td>
<td></td>
</tr>
<tr>
<td>Skill</td>
<td></td>
</tr>
<tr>
<td>• Simulation testing</td>
<td>• Short practice observation</td>
</tr>
<tr>
<td>Attitude</td>
<td></td>
</tr>
<tr>
<td>• Simulation testing</td>
<td>• Long practice observation</td>
</tr>
<tr>
<td>Products</td>
<td></td>
</tr>
<tr>
<td>• Product evaluation</td>
<td>• Product evaluation</td>
</tr>
</tbody>
</table>

An explanation of these groups may be found in Deliverable D2.1.
A next step will be to integrate the findings in the literature review with empirical findings resulting from qualitative study, to be reported in Deliverable 2.2 and 2.3.

5. References
As this is a combination of literature reviews, the references are included in the content section.

6. Tables and Figures
All tables and figures are embedded in the content section.

7. History of the document

7.1 Document history

<table>
<thead>
<tr>
<th>Version</th>
<th>Author(s)</th>
<th>Date</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part 1 V#1</td>
<td>Wisman-Zwarter et al.</td>
<td>December 2014</td>
<td>Deliverance of part 1</td>
</tr>
<tr>
<td>Part 2 V#1</td>
<td>Duijn et al.</td>
<td>December 2014</td>
<td>Deliverance of part 2</td>
</tr>
<tr>
<td>Part 3 v#1</td>
<td>Krull et al.</td>
<td>December 2014</td>
<td>Deliverance of part 3</td>
</tr>
</tbody>
</table>

All versions will be further elaborated, related to Deliverable 2.2 and 2.3

7.2 Internal review history

<table>
<thead>
<tr>
<th>Internal Reviewer</th>
<th>Date</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marieke van der Schaaf</td>
<td>30 December 2014</td>
<td>Some small textual changes</td>
</tr>
<tr>
<td>Sylvia Walter</td>
<td>05.01.2015</td>
<td>Some small textual changes</td>
</tr>
</tbody>
</table>