Assessment of professionalism within dental education
A review of studies

Lina Gassner Kanters

Supervisors: Anders Jönsson and Katarina Wretlind

One-year master thesis in Odontology, Oral health (30 ECTS)
Examensarbete, Odontologie magisterprogram
med inriktning oral hälsa (30hp)
August, 2012
Malmö University
Faculty of Odontology
ABSTRACT

Assessment of professionalism is a concept which has been highlighted during the last ten years. High demands from the public and from authorities combined with an increased educational interest for the concept of professionalism has led to a greater focus on ways of implementing and assessing professionalism in dental, medical and other professional educations. Many authors point out that professionalism is an essential competence, but there is a lack of and a need for a definition of professionalism within dentistry and dental education. It is recognised that professionalism is a broad concept, which calls for different methods of assessments. The aim of this paper is to identify and describe different methods of assessing professionalism within dental education and to categorise them into a blueprint. The studies are discussed out of the perspective of validity and reliability and the methods of assessment are compared to different levels of Miller’s pyramid. Literature is sparse and a mere 16 articles were found to fit the purpose of this paper. Most studies used a traditional way of assessing professionalism, even though research has shown that this way is neither sufficient nor suitable. More research is needed and the methods for assessment need to be further explored. Validity and reliability need to be put in focus, since no method of assessment has proven to be both valid and reliable.
TABLE OF CONTENTS

Terminology 5
Introduction 6
    Assessment 7
    Professionalism – general definitions 8
    Professionalism within medical and dental education 8
    Assessing professionalism 10
        Difficulties with assessing professionalism 10
        The importance of the development of assessments 11
Aim and research questions 12
    Aim 12
    Research questions 12
Material and method 13
Results and analysis 14
    Literature search 14
    Interpretation of literature 17
    Analysis of literature 18
Discussion 25
    Choice of material and method 25
    The result 26
Conclusion 27
References 28
TERMINOLOGY

To make the reading of this paper easier, a short explanation of important terms and concepts is provided.

Assessment: derives from *ad sedere*, to sit down beside, and is primarily supposed to provide guidance and feedback to the learner. An assessment consists of a sample of what students do, in other words gathered information. From this sample, inferences are made, resulting in estimation in the form of a grade, a mark, feedback or recommendations, all depending on the type of assessment. (1, 2)

Formative assessment: an assessment of developmental character, given during a course or a programme and used for training and to provide feedback on the students’ progress. Formative assessment refers to the function of the assessment and not the type of assessment.

Miller’s pyramid: a tool / model that is used within education, in particular creating learning objectives and when working with assessment methods. It was originally developed as a framework for clinical assessment and shows different levels which the student should master when it comes to deliver professional services. The levels are *knows*, *knows how*, *shows how* and *does* (see Figure 1) (3).

Reliability: is a term connected to quality and, simply put, it is consistency and repeatability in measurement. There are different aspects of reliability. In this paper inter-rater reliability is mentioned and this can be tested by letting several persons measure/rate something and then see how equivalent these measurements/ratings are. (1, 2)

Summative assessment: an assessment that is supposed to measure the extent of learning at the end of a course or a programme. Summative assessment refers to the function of the assessment (not the type of assessment). (1, 2)

Validity: Validity, meaning strong, does not have a clear, single definition but generally refers to relevance and whether some kind of data (a conclusion or a measurement) reflects the truth and corresponds accurately to reality. There are many aspects of validity which differs depending, among other things, on the type of research – whether it is qualitative or quantitative. Two common terms is *face validity* and *content validity* – the former being a surface impression of an assessment, the latter requiring the use of subject matter experts and/or statistical tests. As for assessments and validity, the issue is about how plausible the interpretations drawn from test results are, and in what ways the results are being used. (4, 5).
INTRODUCTION

Assessment – a concept indubitably connected to the world of education. In between the curriculum and the objectives, like a big column, stands the assessment. For students often a nerve-racking concept, for teachers associated with hard work and sometimes difficult decisions and, for both, a state of disclosure. Assessments are strenuous, but can be rewarding. Now add the concept of professionalism. Trying to explain that concept is like describing snow in Greenlandic – one will end up with a number of words and views. Assessing professionalism is thus a difficult task and finding good ways to assess professionalism is an even more difficult task. This thesis addresses different ways of assessing professionalism within dental education.

High demands from the public and from authorities combined with an increased educational interest for the concept of professionalism has led to a greater focus on ways of implementing and assessing professionalism in dental, medical and other professional educations (1). There are a couple notions well recognised in literature.

- Professionalism is multi-faceted and difficult to define
- Professionalism should be a core ingredient in professional education
- Assessment is very closely linked with and has an important role in student learning
- Assessments should be valid and reliable and aligned with course objectives
- Assessment of professionalism has proven to be a difficult task due to the complexity of the subject and traditional formats of assessments have shown to be neither sufficient nor suitable (6-10)

To quote Louise Arnold (11), ’Without solid assessment tools, questions about the efficacy of approaches to educating learners about professional behaviour will not be effectively answered’. Lynch et al. (7) states that ’Instead of creating new professionalism assessments, existing assessments should be improved’. Research then need to focus on reform and make better what is already known and done. A validated definition of professionalism in relation to dentistry is a necessity and is called for by Zijlstra-Shaw et al. (10). A definition is of course essential in order for dental educators to be able to teach and assess professionalism and the domains within this construct. Wilkinson et al. (12) have in a review analyzed various definitions of professionalism within medicine and have come up with five main themes, with underlying subthemes. They have also identified assessment tools and clustered them into a tool box, all with the intention of creating a blueprint that matches assessment tools to the identified themes of professionalism. This paper takes the work of Wilkinson et al. one step further by analysing studies of assessment of professionalism within dental education and matching them with the tools and themes of Wilkinson et al.

Based on the Bologna Process and the government bill ‘Ny värld - ny högskola’ (English translation, New world – new higher education), the qualification descriptions in the Swedish Higher Education Ordinance contains learning outcomes which are divided into three groups: knowledge and understanding, skills and abilities, and judgement and approach (13, 14). The third group of learning outcomes has a great deal of importance in, for example, medical, health science, dental and teacher education programs. At the Malmö University Dental Hygiene Education, the learning outcomes within the group judgement and approach aims to develop competences such as empathy, self-knowledge and holistic views and also to present abilities compatible with a professional approach (15). The
faculty of Odontology at Malmö University uses problem-based learning according to the Malmö model (16). The purpose with learning outcomes is that they are to be met by students. To make sure they are met, some kind of assessment is usually needed, but assessing these learning outcomes can be quite a challenge. Several studies underline the difficulties in defining professionalism and finding suitable ways to assess it (7, 11, 12, 17-22).

The main issue is that professionalism is a broad concept with lots of varying definitions that might, and probably will, change over time. Because of this multidimensional nature, there is a need for a number of different ways and tools of assessments (12). Lynch et al. (7) make a note for future research: ‘Instead of creating new professionalism assessments, existing assessments should be improved. Also, more studies on the predictive validity of assessments and their use as part of formative evaluation systems are recommended’. This is supported by other studies which raise the importance of validity in the sense of valid and credible ways of assessment (9, 12, 22). An increased interest in teaching and assessing professionalism is one of the reasons behind the fact that many new assessment tools have been developed, however the same development has not taken place when it comes to finding methods to ensure the quality of these tools. In order to make plausible interpretations from test results, the tools and the scores produced by them need to be validated. There is a lack of research on how to produce evidence supporting this validity, and it is argued that there exists no single method for the reliable and valid evaluation of professional behaviour (11, 23).

For dental educators, an even bigger challenge lies ahead since most research about professionalism has been conducted within medical education (18). Even though the dental and the medical professions differ in many ways, historically and today, we can draw on established thoughts from medicine when working with professionalism within the dental domain (24).

Professionalism within education is seemingly a current global issue, dealt with and written about in many countries as well as Sweden. In Great Britain, the General Dental Council (GDC) expects professionalism to be integrated in dental programmes, with the subject being taught as well as assessed (18). Professionalism is one of seven competencies defined for the Dutch [medical] postgraduate training for family practice (22). In Australia, medical schools are trying to meet professionalism related outcome goals by emphasizing students’ personal and professional development and searching for new methods of assessment (25). Several medical associations in the USA consider professionalism as a very important part of physician competence (9). There are studies that show a connection between unprofessional student behaviour and future, upon qualification, disciplinary actions by medical boards (21).

**Assessment**

Like initially stated, assessment is a vital part of education. Just like pedagogical methods influence learning, assessments do the same. Styles of learning are affected by types of assessments, and students base their studying more on what is being assessed than on what is being taught. The assessment is, just like the objectives, an important part of the curriculum and essential when it comes to promoting learning. There are many different types of assessments and which one to chose depends on which competencies one wants to assess. An assessment gives information about a current status and can be used for developmental and/or judgemental purposes (see formative assessment and summative assessment in the glossary). The assessment can be divided into three parts; taking a sample (the student completes a task), making inferences (an evaluator reviews the completed task), estimating the worth of the
action (the student receives a grade or a rating). Each and single one of these parts can have shortcomings. Validity and reliability should be considered and, if possible, verified for all assessments (1, 2)

Professionalism – general definitions

In the Swedish National Encyclopaedia, professionalism is defined as “professional with a focus on acquisition, one especially in competitive sport used concepts, the opposite is amateurism” (26). In the Oxford dictionary of English, the following definition is found: “the competence or skill expected of a professional: the key to quality and efficiency is professionalism” (27). In the Merriam-Webster dictionary, a professional is amongst other things defined as’ (1): characterized by or conforming to the technical or ethical standards of a profession (2): exhibiting a courteous, conscientious, and generally businesslike manner in the workplace’ (28). These definitions are quite general which means that when terms such as “professionalism” or ”professional approach” are mentioned as learning outcomes, the underlying meaning needs to be clarified.

Professionalism within medical and dental education

In a recently published article, Zijlstra-Shaw et al. (10), like many others before, point out the lack of and need for a definition of professionalism in dentistry. Even though there are more published studies within medicine regarding the concept of professionalism, the same definitional problems exists also here – the essential competence that professionalism is regarded as, has no single, distinct definition. The literature is however filled with recurring themes which has been summarised in a table by Zijlstra et al. (10) and is presented below.

Altruism
Accountability
Autonomy
Compassion
Excellence
Honesty and integrity
Knowledge of ethical standards
Moral reasoning
Reflection
Respect
Self-awareness
Self-motivation, particularly with respect to lifelong learning
Social responsibility
Trustworthiness
Working with others

Several studies have similar descriptions. For example, Veloski et al. (9) have a summary that is almost identical. Arnold (11) mentions several definitions -containing similar terms - of professionalism from North American universities. These two major reviews studying assessment instruments for professionalism has found that professionalism often is evaluated in three different ways: professionalism as a comprehensive construct, professionalism as part of clinical competence and professionalism as separate elements such as empathy, ethical behaviour, altruism and communication. Most studies use the latter description and thus focus on specific elements (9, 11). In this paper, a definition of professionalism by Wilkinson et al.
is used. A study was conducted in 2007-2008 with the aim to match assessment tools to definable elements of professionalism. Various definitions and interpretations of professionalism were classified into five themes and subthemes (see Table 1 below. More on Wilkinson et al. under “Choice of method”) (12).

Table 1. A classification of themes and subthemes, arising from definitions or interpretations of professionalism found in literature published from 1996 to 2007. (12)

<table>
<thead>
<tr>
<th>Theme</th>
<th>Subthemes</th>
</tr>
</thead>
</table>
| Adherence to ethical principles | Including but not restricted to:  
  - Honesty/integrity  
  - Confidentiality  
  - Moral reasoning  
  Respect privileges and codes of conduct |
| Effective interactions with patients and with people who are important to those patients | Including but not restricted to:  
  - Respect for diversity/uniqueness  
  - Politeness/courtesy/patience  
  - Empathy/caring/compassion/rapport  
  - Manner/demeanor  
  - Include patients in decision making  
  - Maintain professional boundaries  
  Balance availability to others with care for oneself |
| Effective interactions with other people working within the health system | Including but not restricted to:  
  - Teamwork  
  - Respect for diversity/uniqueness  
  - Politeness/courtesy/patience  
  - Manner/demeanor  
  - Maintain professional boundaries  
  Balance availability to others with care for oneself |
| Reliability | Including but not restricted to:  
  - Accountability/complete tasks  
  - Punctuality  
  - Take responsibility  
  Organized |
| Commitment to autonomous maintenance and continuous improvement of competence in: | Self. Including but not restricted to:  
  - Reflectiveness, personal awareness, and self-assessment  
  - Seek and respond to feedback. Respond to error. Recognize limits.  
  - Lifelong learning  
  - Deal with uncertainty  
  Others. Including but not restricted to:  
  - Provide feedback/teaching  
  - People management  
  - Leadership  
  Systems. Including but not restricted to:  
  - Advocacy  
  - Seek and respond to results of an audit  
  Advance knowledge |
Assessing professionalism

In order to find the matter of assessing professionalism worthwhile, the fundamental thought must be that people can change. Brown et al. ask if assessing attitudes is worthwhile, and conclude that it is not only worthwhile, but essential. ‘If attitudes cannot be changed, there is little point in assessing them, except perhaps, at the point of entry to dental school or to the profession’ (17). Nash agrees and states that ‘education is a reflective experience that can lead to behavioural change. To suggest that education cannot alter behaviour is to adopt an intolerable cynicism about education’ (29). Due to increased demands from the public as well as from medical, dental and educational associations and organizations, professionalism must have a prominent position in professional education. In order for this matter to be seen as essential also by students, it also needs to be integrated in the curriculum and assessed. Clear expectations and criteria are then needed, which is something that requires a clear definition (1, 9, 10, 21).

Difficulties with assessing professionalism

The difficulties with assessing professionalism have been noted by many educators. Why is it then so difficult? First of all, the definition of professionalism is (clearly) not clear which makes it difficult knowing exactly what it is that is to be assessed. Brown et al. make a point when stating that ‘One needs to have a clear idea of what is being assessed and why’ (17). The reviews mentioned above gives us some notion about the views and research on professionalism but also shows the width of the subject. Different authors, organisations, universities etc. have different definitions - even if there has been development within the medical field on consensus in later years - and professionalism is a complex matter in which one needs to consider many different aspects of an individual’s abilities and competences, and what is important is to make clear definitions before making attempts on creating new assessment methods (12).

Secondly, the difficulties with assessing professionalism are not only due to the lack of clear definitions but also, as mentioned before, due to the multidimensional nature of professionalism which requires an array of different kinds of assessments. Within complex areas there is a need for complex assessment tools and in turn, complex responses require human scorers which in many ways affect the validity and the reliability of the assessment. For example, in a study that aimed to measure the element of humanism, faculty, nurses and patients rated medical residents on three components of humanism – integrity, compassion and respect. To make this assessment reliable, about 50 observations would be needed. Such an assessment is very time-consuming and resource demanding (11, 12).

Another reason that the assessment of professionalism is a challenging one is that professionalism is a subject that differs from many other types of knowledge and skills, this can be exemplified by using Miller’s pyramid (Figure 1, also see “Terminology”). In many cases it can be assumed that if a student on a written test knows for example which forceps to use for which extraction, it is highly likely that the student will use this knowledge also in practice. In this case, the leap between the “knows” level and the “does” level in the pyramid is not a very big one. When it comes to professionalism, proof of knowledge in the lower levels does not necessarily mean that this knowledge will be used in practice. One might know a lot about ethical principles but not know how to use them in practice. The correlation between a correct attitude on a written test and a later displayed behaviour is not always very
Miller’s pyramid can be useful when looking at different types of assessments – in which level do the objectives belong and does this correspond to the assessment method? Like most educators probably know, and Miller states, the top section is the most difficult one to accurately and reliably assess. Due to the close connection between student learning and assessment, Miller suggests that methods for teaching as well as for assessing correspond to the upper section of the triangle (3).

![Miller's pyramid](image)

**Figure 1. Miller’s pyramid, a framework for clinical assessment showing different levels of assessment** (3).

Other difficulties with assessing professionalism mentioned in the literature are; lack of objectivity and personal biases; displayed behaviour might differ from true attitudes and beliefs; reluctance by faculty towards assessing professional behaviour; and findings that perception can be influenced by contextual features (18, 20, 21). Despite the fact that professionalism is in vogue, there is yet to appear a method of assessment that is reliable and valid. In a study by Jha et al. (19), “A systematic review of studies assessing and facilitating attitudes towards professionalism in medicine”, less than half of the reviewed articles demonstrated reliability or validity, and there are just a few instruments that could be said to meet the criteria for an ideal test (10).

**The importance of the development of assessments**

Arnold states that assessment has a central role in promoting professionalism and state of the art examinations are required (11). Fields et al. concludes that methods involved in teaching and assessing professionalism within UK dental schools are largely traditional and that it is necessary for educators to find and implement new and effective methods. A multiple choice test might very well show if a student has specific theoretical knowledge but might on the
other hand not be fitted to assess how a student will behave in practice (18). In other words, work is needed to get to the state of the art examinations Arnold is aiming at.

Why then is it not good enough to use classical knowledge tests when it comes to assessing professionalism? Simply put, because the world is changing and education with it. There is a connection between the work market, ways of teaching and methods of assessments. Today, employees need to be adaptive and possess a range of competencies; they need not only to have factual knowledge and isolated skills. To be a successful participant in the healthcare future one must have basic science and dental/medical knowledge, but of equal importance are communicative skills and ethical awareness and on top of everything, an understanding and interest in life-long learning (24).

What is then known is that professionalism is a central and developing concept in the, amongst others, medical and dental practice and that assessment is needed for the promotion of professionalism. Further work on the definition is needed and elements of professionalism that can be measured needs to be identified. Single assessment tools are not likely to fully cover everything that professionalism is and most likely a combination of tools is required. It is important that these tools are reliable and valid so that credible interpretations can be made from the results (10). Fields et al. note that within the UK, there has been very little investigation of the methods of assessment of professionalism in the dental undergraduate curriculum (18). In several of the above mentioned articles it is concluded that more research is needed when it comes to assessing professionalism in dental education. Yet, there are few studies exploring different methods and it is by far medical literature that dominates the reference lists in featured articles. One could easily say that this is a field where research is lacking – there is a need for dental educators to refine and study methods of assessing professionalism.

**AIM AND RESEARCH QUESTIONS**

**Aim**

The aim of this paper is to identify and describe different methods of assessing professionalism within dental education presented in the literature.

**Research questions**

What kind of different methods of assessing professionalism has been studied and applied within dental education?

How do these studies fit into the blueprint created by Wilkinson et al. (12)?

How do these methods respond to the different levels of the Miller pyramid?

Have validity and reliability been taken under consideration in the studies?
MATERIAL AND METHOD

This is a descriptive, analytical review which was carried out in three stages; 1) literature on assessment of professionalism was studied; 2) a literature search in the database PubMed was conducted to find studies on assessing professionalism in dental education (see Figure 2); 3) studies found in stage 2 were categorized into a modified version of the blueprint by Wilkinson et al. (table 4) and reviewed. The methods of assessment were also linked to a level of the Miller pyramid, and the use of validity and reliability was noted.

For the first stage, reference articles were found through searching the databases CINAHL, ERIC, PsycINFO and PubMed, as well as the database Google. Search terms used were: “assessment”, “assessing”, “professionalism”, “professional attitudes”, “professional approach”, “professional behaviour”, “education”, “dental” and “validity”, in different combinations. The searches were expanded by pearl growing, i.e. by manually checking reference lists. Additional literature was found in libraries. For the second stage, a search was conducted in March 2012 by using the following terms:

<table>
<thead>
<tr>
<th>education OR programme</th>
<th>AND</th>
<th>dental OR dental hygiene</th>
<th>AND</th>
<th>professionalism OR professional behaviour OR professional attitudes</th>
<th>AND</th>
<th>measuring OR evaluating OR assessing</th>
</tr>
</thead>
</table>

The third stage is based on blueprint created in a study by Wilkinson et al. (12). The aim of their study was to define elements of professionalism and match them with tools for assessment. Initially, a literature review of definitions of professionalism was made where upon a thematic analysis followed. The various definitions and interpretations of professionalism were classified into five themes and subthemes (see Table 1). Secondly, a literature review of tools to assess elements of professionalism was made, resulting in the identification of nine clusters of assessment tools. In the end, seven clusters of assessment tools remained since two clusters were considered to overlap others, not being assessment tools in their self. Finally, a blueprint was created where the attributes of professionalism were matched to assessment tools. It was concluded that professionalism is best assessed using a combination of tools and that some existing tools need to be adapted.

The nine clusters, according to Wilkinson et al., are

1) assessments of an observed clinical encounter
2) collated views of co-workers
3) records of incidents of unprofessionalism
4) simulation
5) paper-based test
6) patient opinion
7) self-administered rating scale
8) critical incident report (not in the final blueprint)
9) global view of supervisor (not in the final blueprint)
RESULT AND ANALYSIS

Literature search

The search resulted in 462 publications. After going through the titles, 52 publications remained. The selection was based upon the following criteria

- The publication was written in English or Swedish
- The term “dental” was found either in the title of the article or in the name of the journal
- The term professionalism or attributes connected to the term professionalism, according to a classification of themes and subthemes of professionalism by Wilkinson et al., was found in the title

After reading the abstracts, 18 publications remained. The selection was based upon the following criteria

- The study took part in a dental education setting and included dental/dental hygiene students (as oppose to fully trained practitioners)
- Professionalism or connecting attributes was evaluated, assessed or measured or instruments for evaluating, assess or measure professionalism or connecting attributes were developed

A list of the 18 articles and a flowchart of the selection process follow below.

Number of hits in PubMed and titles to review

\[ n = 462 \]

Excluded publications

\[ n = 410 \]

Number of abstracts to be reviewed

\[ n = 52 \]

Excluded publications

\[ n = 34 \]

Number of full texts to be reviewed

\[ n = 18 \]

Excluded publications

\[ n = 2 \]

Number of articles to be categorized into Table 4 *

\[ n = 16 \]

* Two of the studies, number 3 and 14, could not be categorized into Table 4.

Figure 2. Flowchart, presenting the selection process
Table 3. List of articles, reviewed in full text. Arranged according to year of publication, in descending order.

<table>
<thead>
<tr>
<th>Author/publishing year/reference number</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wener et al. 2011 (40)</td>
<td>Developing new dental communication skills assessment tools by including patients and other stakeholders</td>
</tr>
<tr>
<td>Langille et al. 2010 (39)</td>
<td>The Dental Values Scale: development and validation</td>
</tr>
<tr>
<td>van Luijk et al. 2010 (47)</td>
<td>Promoting professional behaviour in undergraduate medical, dental and veterinary curricula in the Netherlands: evaluation of a joint effort</td>
</tr>
<tr>
<td>Cannick et al. 2007 (33)</td>
<td>Use of the OSCE to evaluate brief communication skills training for dental students</td>
</tr>
<tr>
<td>Christie et al. 2007 (43)</td>
<td>Effectiveness of faculty training to enhance clinical evaluation of student competence in ethical reasoning and professionalism</td>
</tr>
<tr>
<td>Garetto and Senour 2006 (30)</td>
<td>Using an ethics across the curriculum strategy in dental education</td>
</tr>
<tr>
<td>Chamberlain et al. 2005 (44)</td>
<td>Personality as a predictor of professional behavior in dental school: comparisons with dental practitioners</td>
</tr>
<tr>
<td>Kalwitzki 2005 (41)</td>
<td>Self-reported changes in clinical behaviour by undergraduate dental students after video-based teaching in paediatric dentistry</td>
</tr>
<tr>
<td>Zijlstra-Shaw et al. 2005 (42)</td>
<td>Assessment of professional behaviour—a comparison of self-assessment by first year dental students and assessment by staff</td>
</tr>
<tr>
<td>Hottel and Hardigan 2005 (31)</td>
<td>Improvement in the interpersonal communication skills of dental students</td>
</tr>
<tr>
<td>Christie et al. 2003 (45)</td>
<td>Curriculum evaluation of ethical reasoning and professional responsibility</td>
</tr>
<tr>
<td>Chaves 2000 (37)</td>
<td>Assessing ethics and professionalism in dental education</td>
</tr>
<tr>
<td>Röding 2001 (46)</td>
<td>Professional competence in final-year dental undergraduates: assessment of students admitted by individualised selection and through traditional modes</td>
</tr>
<tr>
<td>Author(s) and Year</td>
<td>Title and Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Ryding and Murphy, 1999 (48)</td>
<td>Employing oral examinations (viva voce) in assessing dental students' clinical reasoning skills</td>
</tr>
<tr>
<td>Gleber, 1995 (34)</td>
<td>Interpersonal communications skills for dental hygiene students: a pilot training program</td>
</tr>
<tr>
<td>Bebeau and Thoma, 1994 (38)</td>
<td>The impact of a dental ethics curriculum on moral reasoning</td>
</tr>
<tr>
<td>Bebeau, 1985 (35)</td>
<td>Teaching ethics in dentistry</td>
</tr>
<tr>
<td>Bebeau et al., 1985 (36)</td>
<td>Measuring dental students' ethical sensitivity</td>
</tr>
</tbody>
</table>
**Interpretation of literature**

After reading the publications in full text, 16 out of 18 could be categorised into Table 4. To be able to be categorised, there was a need for the studies to describe a method of assessing professionalism or connecting attributes. The studies were first linked to an assessment tool, and then linked to one or more themes. Two studies were not able to be categorized. One study (47) does not at all mention in what way assessment of professionalism takes places. Another study (48) only assessed clinical reasoning, not in combination with professional communication skills like the abstract indicated. The assessment could not be paired with any of the themes or subthemes defined by Wilkinson et al. (12).

**Table 4. Presenting publications in relation to tools and themes.**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Tool</th>
<th>Assessment of an observed clinical encounter</th>
<th>Collated views of coworkers</th>
<th>Record of incidents of unprofessionalism</th>
<th>Simulation</th>
<th>Paper-based test</th>
<th>Patient opinion</th>
<th>Self-administered rating scale</th>
<th>Critical incident report (not in blueprint)</th>
<th>Global view of supervisor (not in blueprint)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>30, 31</td>
<td></td>
<td></td>
<td>37, 35, 36</td>
<td></td>
<td>37, 30, 35, 36</td>
<td>40</td>
<td>40, 41</td>
<td>43, 44, 45</td>
<td></td>
</tr>
<tr>
<td>#2</td>
<td>31</td>
<td></td>
<td></td>
<td>33, 34</td>
<td></td>
<td>34, 39</td>
<td>40</td>
<td>40, 41</td>
<td>44, 46</td>
<td></td>
</tr>
<tr>
<td>#3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>39</td>
<td></td>
<td>40</td>
<td>40, 42</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>42</td>
<td></td>
<td>42, 44, 45, 46</td>
<td></td>
</tr>
<tr>
<td>#5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40</td>
<td>40, 41, 42</td>
<td>42, 44, 46</td>
<td></td>
</tr>
</tbody>
</table>

**Specification of themes**

#1 Adherence to ethical principles
#2 Effective interactions with patients and with people who are important to those patients
#3 Effective interactions with other people working within the health system
#4 Reliability
#5 Commitment to autonomous maintenance and continuous improvement of competence

For further information on what is included in the different themes, see Table 1

The tool most used was “paper-based test”, followed by “simulation” and “global view of supervisor”, the latter tool does however not hold a place in the final blueprint by Wilkinson et al. (12). Three of the nine tools were not used at all. The themes assessed most often were #1 and #2, “adherence to ethical principles” and “effective interactions with patients and with people who are important to those patients”. The least assessed theme was #3, “effective interactions with other people working within the health system”. 
Analysis of literature

Below, a short description of each tool is provided and the publications linked to the tool are described and discussed.

Assessment of an observed clinical encounter

This type of assessment can, according to Kramer et al. and Wilkinson et al. (2, 12), be good for evaluating specific areas of performance related to clinical competencies, e.g. behaviour, critical thinking and interpersonal skills. Wilkinson et al. mentions a couple different tools, such as the Standardised Direct Observation assessment Tool, a 26-item checklist used to evaluate core competencies by direct observation in medical education, and the Mini-Clinical Evaluation Exercise (mini-CEX). The mini-CEX is a test where a trainee interacts with a patient, conducts a physical examination, provides a diagnosis and a treatment plan, all whilst being observed by a single faculty member. A structured document is used for scoring and the encounter is supposed to be relatively short, about 15 minutes. Depending on surrounding factors such as the rater or specificity of the checklist, this method might be susceptible to personal biases and subjectivity (2, 12).

Two of the articles (30, 31) found in the search can be linked to this type of assessment. Garetto and Senour (30) do not in detail go into the assessment but the students are assessed on professionalism by attending clinic faculty for each clinical encounter with a patient. Clinic Directors independently assess each student at the end of each academic year. It is not clear for how long each encounter is observed, or if one or more faculty members perform the assessment, which makes for a discussion whether this should be classified as an “assessment of an observed clinical encounter” or rather categorized into “global view of supervisor”. At the Indiana University School of Dentistry, two other types of assessment are also used; one of them being written essays has been categorized into “paper-based test”. The third type if assessment is a Triple Jump Examination (TJE) which primarily is an oral assessment, but also has other elements. This type of examination is often used within problem-based learning and was developed at the McMaster University in Canada. The TJE has not been found to match any of Wilkinson’s tools. Interestingly, all three assessments have a pass/fail grade system. In the authentic assessment mentioned above, a passing mark on continued competency in professional responsibility is required irrespective of technical performance. The ethics and professionalism part of the curriculum is quite ambitious, consisting of 37 problem-based learning cases, as well as several other activities. It is however not clear what exactly is included in the term professionalism and the emphasis lies on ethics, whereby the article is placed under theme #1, adherence to ethical practice principles, in table 4. Reliability and validity was not discussed in the study.

The second article, by Hottel and Cardigan (31), describes the assessment more detailed. The article is considered to cover theme #1 and #2 in table 4. The aim of the study was to study, evaluate, and assess the outcomes of a course designed to help dental students develop interpersonal communication skills. The students were evaluated interacting with patients, before and after a 35-hour training course, by psychology graduate students, with the use of a behavioural observation form. The form had a five-point, ten-item scale and measured expressive and receptive communication skills, nonverbal communication skills, professional presentation, and sensitivity to cultural, ethnic and gender diversity. The rating scale has been used for the same purpose in two previous studies and was then found to be a reliable measure. In this study, internal reliability of the rating scale was ensured by calculating
Cronbach’s Alpha, resulting in a value of .74, and content validity was considered ensured by ‘carefully evaluating the full range of behaviours to be observed’. Inter-rater reliability was also established. Results show statistically significant improvement in all ten items, proving that the course was effective.

The second study displays an interdisciplinary arrangement, connecting psychology students and dental students. That type of assessment could, just like the one presented in the first study, be used with a pass/fail system, e.g. by using minimum scores. Another way of using it is by checking improvement over time, which also gives better reliability. A couple of things need to be taken under consideration though. First of all the cost and expenditure of time of the quite extensive course, and second of all, the fact that the assessment does not cover the whole range of themes and therefore need to be complemented by other assessments. As for this tool, “assessment of an observed clinical encounter”, an authentic interaction is good for validity aspects. To achieve reliability, potential error and bias need to be minimized since observations are subjective in nature. Regarding Miller’s pyramid, this tool corresponds very well to the top section, to the level “shows how”.

**Collated views of co-workers and Record of incidents of unprofessionalism**

No articles were found to correspond to these two tools. A collated view of co-workers is a sort of multisource feedback, sometimes referred to as a 360-degree assessment. Within medical education there are few published data on the outcome of multisource feedback, which can be used to assess competencies difficult to assess within formal assessments. It is most effective when the sources, for example peers, are credible and the feedback given is constructive. Criteria need to be clear and the content needs to be relevant and be connected to essential knowledge and abilities. Peer assessment might be favourable to use when students already contribute to and are used to group learning, like in problem-based learning (2, 32).

Record of incidents of unprofessionalism could be used by an incident reporting form; the form would be looked at by a group who would decide if further action was to be taken. Like mentioned above, no articles regarding dental education were found, but at University of Queensland, Australia, students at the medical education have been assessed by so-called needs assistance reports. The reports are part of a personal and professional development (PPD) process with a primary function of supporting referred students but also include an assessment function for unprofessional attitudes and behaviours. From 2000 to 2006, two students failed a year on the basis of unprofessional behaviour, and so this assessment has been used with a pass/fail system (21).

These methods could initially be thought to fit right in to the “shows how” level in Miller’s pyramid, but a limited direct observation does not really set the standard for a reliable measure. Multisource feedback could however quite well be a type of assessment that could match a student’s true performance (3).

**Simulation**

Simulations are thought-out scenarios, resembling real-life situations using models or simulated patients. An objective structured clinical examination (OSCE) is one way to use simulation, and can advantageously assess infrequent or unpredictable situations. It can be used for assessing how someone handles pressure, and/or communication skills. Validity can be reduced due to the artificial context, and it is therefore important that the scenarios or
situations are as realistic as possible. Just using one single OSCE station can be judged as unreliable. If the assessment is accounted for as valid and reliable, it fits right into the upper level of Miller’s pyramid, the “shows how” section. An unpredictable situation, in which someone is exposed to pressure, could end up in the highest section since it shows a true action – it is however still most likely an artificial setting which is something one needs to bear in mind (2, 12).

Five articles have been found to use or describe simulation (33-37). The first one, by Cannick et al. (33), is considered to cover theme #2 in table 4. The aim of this study was to evaluate dental students’ competency in interpersonal and tobacco cessation communication skills after a two-hour training session - not to directly assess professionalism. However, communication is a vital part of professionalism and embedded in several of the Wilkinson themes. In the study, the students are evaluated on two OSCE stations; one regarding patient/dental student interaction and one regarding tobacco cessation. On the former one, the students were rated on a scale from 1 (poor) to 5 (excellent) on six checklist items. At the latter station, either a “1” or a “0” was given, depending on if the student performed the item or not (e.g. “Asked about tobacco use”). The study showed that the two-hour training session had no statistically significant effect on the results of the OSCE, however all first-year students (intervention as well as control group) increased in tobacco cessation communication scores. The authors suggests ‘that a comprehensive communication skills training course may be more beneficial than a single, brief training session for improving dental students’ communication skills’.

Reliability and validity was not discussed in the study, but standardisation was ensured by giving evaluators checklists of the criteria to be covered. Since both stations assess communication skills, this counts for better reliability. Students were videotaped at the tobacco cessation station and the tapes were viewed and evaluated by two reviewers which is also an advantage (compared to using only one reviewer). The standardized patient used a checklist to evaluate the students. It is important that the checklist criteria are clear and that the same patient is used for all students and both test rounds - if several patients are used they should be calibrated.

The second article, by Gleber (34), also covers theme #2 and “simulation” as well as “paper-based test”. The study aims at determining if the Carkhuff model of communication skills training would improve the interpersonal communication skills of junior dental hygiene students. One experimental group of dental hygiene students received 20 hours of training. Data were collected pretest, posttest and one year later using three instruments. Results show significant differences in knowledge but not in behaviour. After one year there was significant improvement in both knowledge and behaviour. Note that change is recorded – there is no pass/fail grade. The three instruments are
1. Interviewing a simulated patient -“simulation”
2. Written responses to brief interaction scenarios -“paper-based test”
3. Mehrabian and Epstein measure of emotional empathy -“paper-based test”

In the simulation part of the study (instrument 1), behaviour is assessed, whereas for the paper-based tests, knowledge is assessed. Each interview lasted for ten minutes and each student had an interaction with a patient who presented a personal dental hygiene problem. Video recordings were rated by three raters, however data from one rater was eliminated, using Carkhuff’s levels of helping. Inter-rater reliability was determined by Chronbach’s alpha. No difference between the groups was shown post-training, but positively both groups increased mean scores. At the one-year follow up there was a statistically significant difference between the groups, with the experimental group increasing more on the five-point
Levels of Helping Scale. The study shows that the students improved more regarding knowledge than regarding behaviour. Since “knows”, or “knows how”, forms the base of the Miller pyramid it might be preferred to give the students training that increase their behaviour as much as their knowledge.

Two of the remaining articles were, as well as another article using a “paper-based test”, written by the same author, Bebeau (35, 36). The author is focussed on ethics whereby the articles end up belonging to theme #1, adherence to ethical principles. The articles describes a curriculum project regarding ethics at the Minnesota dental faculty and three assessments connected to this; the dental ethical sensitivity test (DEST), which is classified under “simulation” due to the fact that it’s an audio taped test, the Defining Issues test (DIT) and an essay, the two latter tests being classified as “paper-based tests”. The development of the DEST is also described. Four dramas were created for the DEST. The dramas were based on dentists’ reports of frequently occurring ethical problems and issues. Dentists, specialists and one psychiatrist cooperated in the development of the dramas that were later audio taped and tested extensively by student groups. Students listened to the dramas and were asked to pretend to be the dentist in the drama. At a certain point in the dialogue they took the place of the dentist on the tape and carried on the dialogue, following this they answered 8 questions. They were scored on several items for each drama and also given a total score/drama (scale 1-3). In total there are 34 ratings distributed across the four dramas. Tape recorded responses were used to establish inter-rater reliability. Test-retest reliability is estimated to be good but the result is tentative. The DEST requires a lot of work in the initial, developing phase however less work once criteria are set. The author suggests peer-review as an alternative. No pass/fail grade is used for the DEST in this study but the test could be used for checking development regarding ethical sensitivity.

The fifth article, by Chaves (37), describes how ethics and professionalism is assessed at the Indiana University School of Dentistry. Here, the DEST and the DIT are also used. Like briefly mentioned above about tentative results for reliability, Chaves also notes that information regarding validity and reliability for the DEST is sparse due to the fact that it is a test intended specifically for the dental profession, and not previously used in, for example, medicine. Chaves also describes the DEST as a formative as well as a summative tool. The DEST does not qualify itself to the upper section of Miller’s pyramid, but corresponds to a students’ “knows how” level.

**Paper-based test**

Merely from the name, it is quite obvious that the tests using this tool end up rather low in the Miller pyramid, in the “knows” or “knows how” sections. A paper-based test could for example be a multiple-choice test or the DIT. It can assess what someone knows about principles for ethics and professionalism, moral reasoning, or in some cases problem solving, but it cannot assess behaviours or what someone would do in reality. When it comes to multiple-choice tests, they have a low level of fidelity to actual practice and are susceptible to measurement disturbances such as guessing or hedging. A structured essay has the strength of being able to assess most competencies. The students are evaluated on their ability to clearly identify a problem and come up with a logic solution. The evaluation might be prone to be subjective and also influenced by distractions such as spelling mistakes and grammatical errors (2, 12).
Seven articles have been found using paper-based tests (30, 34-39). Three of them are by the same author, Bebeau, describing the Defining Issues test and an essay (35, 36, 38). The article by Chaves (37) also describes the DIT. These articles covers theme #1, adherence to ethical principles. The DIT is a well proven multiple-choice test derived from Kohlberg’s descriptions of stages of moral judgement development. The test consists of six moral/ethical dilemmas. The students are to rate the importance of twelve statements that describe important considerations and are designed to reflect different stages of ethical development. Reliability and validity for the DIT, which has been used in more than a thousand studies, are well established.

The fifth, and yet another article covering theme #1, is a descriptive one, written by Garetto and Senour (30) who describe the curriculum in ethics and professionalism at the Indiana University School of Dentistry. This article is mentioned above, in “assessment of an observed clinical encounter”. The paper-based test has the form of written essays discussing clinical scenarios. The students’ ability to provide responses using ethical principles and value arguments is evaluated and the students must attain a “meets expectation” score on their essays. If failing, it is necessary to rewrite the essay addressing specifically directed critique. Reliability and validity was not discussed in the study.

The article by Gleber (34) mentioned above under “simulation”, covers theme #2. Three instruments are used; two of these are paper-based tests. The written responses to several brief interaction scenarios were evaluated by three raters. Five suggested responses that correlate with Carkhuff’s Levels of Helping are given; the students are to evaluate these according to a Likert-type scale. Inter-rater reliability was determined by Chronbach’s alpha. The third test was the Mehrabian and Epstein measure of emotional empathy. The students are to give a response to 33 short statements. An example of a statement is: “it makes me sad to see a lonely stranger in a group”. No significant change in emotional empathy scores could be attributed to the interpersonal skills training in this study. In addition, emotional empathy was demonstrated to have no predictive value as a determinant of student knowledge or behaviour of interpersonal skills.

The last article in this group is an article by Langille et al. (39) describing the development of the Dental Values Scale (DVS). The DVS is a twenty-five-item, five-factor model which the authors consider to be a reliable measurement tool because of good internal consistency values. The five factors are altruism, personal satisfaction, conscientiousness, quality of life and professional status which link this article to theme #2, #4 and #5. The DVS cannot be used with a pass/fail grade but can function as a training tool with a connecting development programme. Since the DVS measures values it places itself very low on the Miller pyramid, in the “knows” section.

Patient opinion

It is difficult to achieve good reliability when it comes to patients’ ratings – according to Epstein (32), one might need as many as 50 patient surveys. Patients’ ratings are also typically high, although some patient populations can be more critical than others, and patients might not be able to distinguish different elements of clinical practice or understand the language used in the survey. Patients’ opinions should be used in addition to other assessments. Just like “collated views of co-workers”, it can be good when assessing competencies such as communication and interpersonal skills that are difficult to assess using formal assessment conditions. Typically, these surveys consist of rating scales or checklists where patients are
asked to declare their satisfaction with the performance of the student, usually using
categories such as poor, fair, good and excellent (or corresponding). It has to be considered
that the patients need to be able to complete the survey in a reasonable amount of time, about
ten to fifteen minutes (2, 12, 32)

One article, by Wener et al. (40), was found matching this tool, it also matches the following
tool, “self administered rating scale”. The aim of the study was to produce an instrument to
assess students’ clinical communication skills. Just like others before, the authors recognize
that it is a difficult task to teach and assess communication skills using traditional didactic
methods and written exams. A key point in the study was that ‘evaluating effective
communication should not only be based on the observable behaviours of the health
professionals, but also on the behaviours and perception of the patients’. One instrument was
developed for patients, the Patient Communication Assessment Instrument (PCAI), and one
for students, the Student Communication Assessment Instrument (SCAI). Both dental and
dental hygiene students, as well as patients and other stakeholders such as clinical instructors
and support staff, were involved in the development of the instruments. The final instruments
are identical, except for grammatical alterations and seven demographic items that are only
applied to the patients. The instrument consists of seventy-two items divided into seven
categories, which are: telephone; initial greeting; relationship building, trust, and respect; non-
verbal communication; sharing information and decision making; attention to your comfort;
and team communication. Validity and reliability is yet to be determined. Themes covered by
this study are #2, #3 and #5. In regards to the Miller pyramid, this type of assessment reaches
the upper section, the “shows how” level.

Self-administered rating scale

According to Wilkinson et al., this tool is questionnaire-based and can aid reflection and
assess personal attributes and attitudes but it is difficult to assess actual behaviour (12).
Kramer et al. suggests that this type of assessment can be applied to most competencies,
underlying knowledge or ability except for highly precise technical skills. It is best if there are
clear criteria to judge yourself against and known standards for acceptable levels of
knowledge and performance. Self-assessment works well in a preclinical setting and can be an
important part of formative assessment. It can however be limited due to a lack in self-
awareness (2). As for the Miller pyramid, this type of assessment can be placed in the “knows
how” or the “shows how” level, depending on the content of the questionnaire and the follow-
up.

Three articles (40-42) were found matching this tool, one by Wener et al. (40) is mentioned
above and had the aim to produce an instrument to assess students’ clinical communication
skills. Questionnaires were filled out by the patients as well as by the students. The students
filled out the SCAI by answering questions like “how were you at treating the patient/client
with respect”, it is however not clear to what extend the students received feedback or if
there, in any way, were any requirements regarding the result.

The second article, by Kalwitzki (41), aims at evaluating the aspects of clinical behaviour that
the students felt had changed after video-based teaching stressing especially the practitioner-
patient interaction. Students were video-taped, and then chose a 10-minute sequence that was
shown to and discussed in the student group. They then filled out a questionnaire consisting of
36 questions regarding behavioural changes, focusing on patient management, infection
control, and ergonomics. The main focus of the study was if seeing themselves on video
would change the students’ behaviour. However, by combining this with a questionnaire the
students can become more aware of their behaviour and this can function as a formative
assessment, alternatively as a training tool for professional behaviour. The questions on the
questionnaire could be modified to fit the purpose. Reliability and validity was not discussed
in the study, which covers theme #2 and #5.

The third article, by Zijlstra-Shaw et al. (42) had the aim to assess usefulness and
acceptability of a method of assessing professional behaviour of undergraduate dental
students. A form, with an ordinal scale, was developed with three main criteria – carrying out
assignments, behaviour in the pre-clinic, reflection upon own performance. There were also
supplement criteria. The criteria were former used in medical faculties and were adapted to fit
dental students. The answers were marked on a five-point scale as well as with positive and
negative “smileys”. The forms, assessing professional behaviour, were filled out both by
students (self-assessment) and by staff (assessing the students). The assessments were
compared and the students then, during a meeting, received feedback from the staff. In
general, the students gave themselves more positive smileys than the staff, and the mean score
from the main criteria was lower from the staff for all three criteria. In the study, the response
rate was high, but it is important to remember that enough time needs to be set aside for filling
out the forms. It is concluded that ‘furthermore research need to be done on the validity and
reliability of the forms if they are to be used to produce a measurable outcome’. The study
could, with some modification like for example introducing minimum scores, quite likely be
used for assessment with a pass/fail grade. The study covers theme #3, #4 and #5.

Critical incident report (not in blueprint)

No studies matching this tool were found. The method asks the student to reflect upon a
critical incident that has been witnessed or experienced. It is considered to contrast with the
tool “record of incidents of unprofessionalism” and this tool is not a part of the final blueprint
since the areas it maps against would be idiosyncratic and individual.

Global view of supervisor (not in blueprint)

Even though this tool not ended up in the final blueprint, quite a few studies (five) were found
using this type of assessment (42-46). One of them has already been discussed under “self-
administered rating scale” (42). The method is a summary view reported on a form with
predefined criteria. Since only a single observer is used, the method is considered to be
unreliable. It can however be useful if it is combined with other assessments and used
repeatedly. If several raters are used, this functions as multisource feedback and thus
correlates with the tool “collated views of co-workers”. Of course, feedback from supervisors
is something essential, but not considered to be fitted for the final blueprint in the study by
Wilkinson et al. In regards to the Miller pyramid, this type of assessment reaches the upper
section, the “shows how” level.

A study by Christie et al. (43) study focuses on authentic evaluation, connected to core values
(the ADHA code of ethics). Focus in the article is not on the process of the assessment but on
how the faculty can become better at evaluating the students and connecting their feedback to
core values, which also is an important research field within this context. The study represents
theme #1.
Another study, “Personality as a predictor of professional behaviour in dental school: comparisons with dental practitioners” (44), aims at developing a measure of performance that would assess the student’s degree of professionalism. The performance measure is based on the seven Canadian Dental association competencies which are used in admissions procedures in Canada. For each competency, a scale value of 1, 3 or 5 was created by focus groups. Each student was evaluated once. The test has a high degree of internal reliability. Something to consider here is that if attitudes and behaviours are adaptable and can develop and change, admissions tests are not what should be focussed on. Instead time should be spent on the curriculum and the assessments during the education programme (17). One positive thing with the study is that it covers four themes out of five, all except for theme #3.

In another study by Christie et al. (45), the assessment process is not in detailed described but the article mentions several different ways of assessing whereof one could be categorized into this tool. The students are evaluated on case management for each patient treated, by a clinical faculty member. The evaluation is based on a five-point rating scale and there’s the possibility of adding a comment. The comments were analyzed to determine if they were related to core values and ethical principles taught in the curriculum.

The main aim of the fifth and last article in this tool-group is not to assess professional competence but to evaluate the mode of admission. The assessment of professional competence is however related to the mode of admission. Professional competence was assessed at one occasion by faculty members who knew the students well and had followed the students from one to three semesters. The assessment was based on seven different predefined criteria, to be graded on a scale from 1 (improvement essential) - 5 (excellent), and one overall rating. The seven criteria were knowledge, initiative, responsibility and judgement, patient contact, clinical skills, co-operative approach and commitment and motivation. Validity and reliability was not discussed in the study (46).

Exceptions

Two articles were not at all able to categorize into Table 4 (47, 48). The first study is a qualitative survey that describes the outcome of promoting teaching and assessment of professional behaviour in medical, dental and veterinary education in the Netherlands. Although behavioural assessments take place at several programmes across the Netherlands, it is not in any way described how (47). The second article focuses on clinical reasoning skills and not on communication as written in the abstract. The aim of the study is to develop and evaluate a comprehensive oral examination system and the article could be very useful for dental educators wanting to improve an assessment, it does not however fit the purpose of this thesis (48).

DISCUSSION

Choice of material and method

The article by Wilkinson et al. was chosen as a base because of several reasons. It is a fairly new (2009) article, it builds on extensive reviews - amongst others a thorough literature review by Veloski et al. on assessments instruments for measuring professionalism - also used as references in this paper, and it summarizes both professionalism and methods of assessments in a comprehensive way. Although it might be a disadvantage that the article
deals with professionalism within medicine and not dentistry, there is no corresponding article, which gives such a thorough analysis, in dentistry.

Miller’s pyramid was used to emphasise the complex and the multidimensional nature of the construct of professionalism. The way in which the model could relate to a variety of competencies was considered useful. Although the ultimate goal is that dental practitioners will succeed on the top level, “does”, by providing excellent care, all the levels in the pyramid are essential and can be used for different competencies and different assessments. Parallel’s can be drawn between Miller’s pyramid and the three groups of learning outcomes described in the Swedish Higher Education Ordinance. Just like the complexity increases upwards in Miller’s pyramid, the three groups of learning outcomes can be considered to increase in complexity. The base would be formed by from “knowledge and understanding”, the middle section by “skills and abilities” and the top section by “judgement and approach”. There are of course other models that could have been used, e.g. Bloom’s taxonomy or the SOLO taxonomy, but Miller’s pyramid was considered to exemplify professional development in a perspicuous way (1).

Validity and reliability are two components that are essential to keep in mind when discussing assessment. The choice for this paper was to explore if validity and reliability was something that had been taken under consideration at all in the studies, not to go deeper into analysing what methods or approaches that had been used. A deeper analysis of validity arguments for different assessments of professionalism could be a useful complement. One could for example carry through a structured argument on validity using Michael Kane’s framework, however that is a different paper (23).

One could discuss the decision to only use one database to find studies to match the aim, however literature on professionalism has been thoroughly reviewed and no further studies have been found or been referenced to in any articles. Even by checking the reference list of the latest publication on the subject, no further studies were found. Out of 112 references in the article by Zijlstra-Shaw, a mere 25 are related to dentistry or dental education which displays the lack of a dental perspective (10).

The result

The aim of this paper was to identify and describe different methods of assessing professionalism within dental education. Literature in this field is sparse and only 16 studies fit the inclusion criteria. The studies were categorised into a table by being linked to an assessment tool as well as one or more themes, including subthemes, of professionalism. It was found that five out of seven tools part of the blueprint by Wilkinson et al. had been used in the studies. On top of that, five studies had been conducted using a tool (“global view of supervisor”) that did not fit into the final blueprint. No studies were found to match the tools “collated views of co-workers” and “record of incidents of unprofessionalism”. Seven studies had used the tool “paper-based test”. This is interesting, considering the knowledge that traditional formats of assessments have shown to be neither sufficient nor suitable, but hardly surprising. A paper-based test is probably the type of assessment that is the easiest to manage, the least time-consuming and it is a traditional method that most people are familiar with. It is a suitable method if the objectives for the assessment are relating to the lower section of Miller’s pyramid, and for assessing theoretical knowledge. A possible way to use a paper-based test would be during the initial phase of an educational programme, e.g. during the first
semester, to make sure that the students have knowledge of ethical principles and other basic concepts.

Five studies were found to have used the tool “simulation”, a tool which has the advantage of making students show how they do something. The negative thing is that the five studies only represent theme #1 and #2. Since simulation can be used to assess more complex objectives, it would be a good idea for the future to study how simulation can be used in a wider sense, covering also the other themes in the blueprint.

Three studies used the tool “self-administered rating scale”. This method fits very well with a problem-based learning curriculum. Although in most cases not suitable for summative assessments (article number 9, if modified, could be an exception), this method works fine for formative assessments and can be used to study student development. An assessment of this type could also form the base for appraisals or development discussions.

Two studies used the tool “assessment of an observed clinical encounter”. They cover only theme #1 and #2, but this type of assessment could very well be suited for assessing the other themes as well. Due to being authentic, validity is good and by checking surrounding factors such as a specific checklist and objectivity, good reliability could also be reached. Since this type of assessment reaches the higher levels in Miller’s pyramid, a development of existing methods, e.g. the study by Hottel and Hardigan (31), would be a good research area.

One study used the tool “patient opinion”, and in the same study “self-administered rating scale” was also used. Since the public are who we serve, including the patients in assessments is interesting. Patients are of course often involved in studies, but obviously not very often involved in the assessment process since only one study was found. Although reliability is an issue here, questionnaires from patients could function as formative assessments and, like a “self-administered rating scale” be used for clinical development discussions.

Reliability and validity has been discussed unevenly throughout the studies. In some studies it is not mentioned at all, whereas in others it is briefly addressed, usually by concluding that it is lacking and more research is needed to ensure reliability and validity. Inter-rater reliability is by far most mentioned. The Defining Issues Test is the only assessment with well established reliability and validity. When constructing a new assessment, reliability and validity should a part of the process at an early stage.

**CONCLUSION**

The result in this paper is limited since there are only 16 studies conducted during 26 years. Most studies used a traditional way of assessing professionalism, even though research has shown that this way is neither sufficient nor suitable. The studies included in this paper have been identified, described and analysed which has contributed to a deeper understanding than if the studies had only been described. The work by Wilkinson et al. has functioned as a mirror, reflecting different approaches to the concept of professionalism as well as to methods of assessment. The blueprint that Wilkinson et al. has created could be a starting point for a blueprint with a dental perspective. During an educational programme, professionalism should be assessed at several occasions by using different methods, covering all levels of the Miller pyramid. A lot of articles have been written about the need for a definition and the need for assessment of professionalism, however there has been little progress despite this consensus. More studies are needed and all aspects of professionalism need to be taken under
consideration. Further investigation of the validity and reliability aspect would make the work with assessing professionalism in a good way – not just assess it – easier. This paper is an attempt to illustrate just how complex the concept of professionalism is and how difficult it is to assess. It is also an attempt to highlight the need for more studies and a need for progress regarding a consensus on a definition of professionalism within dentistry.

REFERENCES


28


