



ORIGINAL ARTICLE

# Reliability of rubrics in the assessment of orthodontic oral presentation



Naif A. Bindayel

Department of Pediatric Dentistry and Orthodontics, College of Dentistry, King Saud University, P.O. Box 60169, Riyadh 11545, Saudi Arabia

Received 14 June 2017; revised 23 July 2017; accepted 24 July 2017  
Available online 2 August 2017

## KEYWORDS

Assessment;  
Education;  
Orthodontic;  
Oral presentation;  
Reliability;  
Rubrics

**Abstract** *Aims:* The aim of this study was to evaluate the reliability of using rubrics in dental education, specifically for undergraduate students' assessment in orthodontic oral presentation.

*Methods:* A rubric-based case presentation assessment form was introduced to three contributing instructors. In each instructor's group, the course director, along with the assigned instructor, assessed 8 randomly selected fourth year male dental students utilizing the same assessment form (total of 24 students). The two final scorings made by the assigned instructor and the course director were then gathered for each student. The data of this prospective comparative study then was analyzed using paired *t*-test to look for any significant differences in the scoring of the course director and each instructor in each group.

*Results:* No significant statistical differences were detected in grading variables between the instructors and the course director. Furthermore, the data showed no significant correlations between the students' final course grade, and their case presentation grades scored by instructors'/ course director.

*Conclusion:* Despite the elaborate nature of the routine orthodontic case presentation, the use of rubrics was found to be a promising reliable assessment element.

© 2017 The Author. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Teaching and healthcare practice are interrelated. This is due to the service delivery system that requires the attendance of different personal with different levels of knowledge and expe-

rience. Teaching in the clinical environment is defined as teaching and learning focused on, and usually directly involving, patients and their problems (Spencer, 2003). And it is interesting to know that the word 'doctor' is derived originally from the Latin word "docere", which means "to teach" (Shapiro, 2001).

Whether in healthcare profession teaching or not, the process of learning and student comprehension is complicated. Many methods have classically proposed ways of thinking behaviors that is believed to be important to the process of learning. Bloom's taxonomy was among the earliest and focused on the knowledge (cognitive) domain (Bloom and

E-mail address: [nbindayel@ksu.edu.sa](mailto:nbindayel@ksu.edu.sa)

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

Krathwohl, 1956). Other domains focused on the attitude (affective) domain (Krathwohl and Bloom, 1964) and skills (psychomotor) domain (Simpson, 1972). Curry Onion-Model of learning described further the different aspects (i.e. as layers of onion) of learner and how they learn (Curry, 1983). Each style is characterized by specific features including the ability to acquire knowledge, sort and store information, learners' interaction with peers and society. Making an assessment to test the learner should touch and consider these styles.

Based on the various domains incorporated into the leaning system, an ideal process of student assessment should cover the attitude, skills and knowledge domains. This can be a complex task, however, the awareness of the importance of these aspects in the assessment process is essential.

Assessment can be formative or summative. Formative assessment is essential for monitoring performance during a program of study, while summative assessment usually done at the end of a program such as competency and licensing examinations. Whether formative or summative, methods of assessment vary and require critical planning where any chosen method of assessment must reflect on the nature of the acquired knowledge being tested.

Many evaluation models were proposed based on each learning domain. The objectives approach (Tyler, 1949) provide a consistency between goals, experience, and outcomes. It includes pretest and posttest design that students' progress can be measured from. The Goal-Free Assessment model (Scriven, 1991) advocates the implementation of an external evaluator whom is unaware of the stated goals and objectives. The value of a program will be determined based on the outcomes of a program and its quality. Unlike the CIPP model (context, input, process, and product), where the information for assessment is being gathered from a variety of sources to provide basis for making better decisions (Stufflebeam, 2003). Other models were also proposed such as the Hierarchy of Evaluation model (Kirkpatrick, 1979), and the Naturalistic model (Guba, 1978). Additional assessment method that was found to be reliable in clinical setting for health care professional is the RIME method (Pangaro, 1999). It has four stages of students' development beginning with being a reporter, interpreter, manager, and then educator that leads to professionalism in medicine.

The assessment is the curriculum, as far as the students are concerned (Ramsden, 1992). And whether or not any these assessment models are being adopted, the assessment process has to be undertaken properly to be reflective of the actual students' actual learning. One of the tools used in assessment nowadays is the use of rubrics. Rubrics can be defined as: a scoring guide or scale consisting of a set of criteria that describe what expectations are being assessed/evaluated and descriptions of levels of quality used to evaluate students work or to guide students to desired performance levels.

The use of rubrics has many advantages such as enhancing the quality of direct instructions, save the time used for explaining the assignment, and increases the efficiency of marking (Hancock and Brundage, 2010), and produce grading calibration (Turbow et al., 2016). It improves the quality of students' projects outcomes by providing clear guidelines regarding the expected criteria. It simply fulfills the required need of shifting the assessment methods from being subjective, to fairly objective.

Rubrics are mainly of two types, analytic and holistic. The analytic type is a more of detailed version of rubrics that identifies and assesses the individual components of a completed project. While the holistic assesses student work as a whole. There are also some subtypes of rubrics such as weighting rubrics. Weighting rubrics is an analytic rubric in which certain concepts are judged more heavily than others (Dong et al., 2011).

The process of formulating rubrics can be initially difficult; thus, it requires support, time, and practice. It mainly consists of three major steps. First, the evaluation criteria and the concept being taught have to be chosen. This step is followed by organizing these criteria, and developing a grid and inserting criteria.

In the last decade, rubrics were incorporated in the teaching curriculum of many fields. Recent literature shows its wide applicability and acceptance in the teaching of medicine (Baldwin et al., 2009; D'Antoni et al., 2009), nursing (Daggett, 2008), and pharmacy (Blommel and Abate, 2007). In dentistry, Assessment rubric was used for third year dental student in developing a course toward mastering sound communication skills with patients (White et al., 2008). Also, scoring rubric was implemented to evaluate dental student portfolios as a mean of student competency assessment (Gadbury-Amyot et al., 2003).

Oral case presentation typically included in most of health-care taught courses. The task allows for initiation of self-learning process, and assess clinical reasoning competency (Wiese et al., 2002), thus requires a crucial assessment tools to reflect student's comprehension. Peer assessment is widely used in this field as an effective formative assessment tool (Speyer et al., 2011). Other methods including the use of rating scale (Lewin et al., 2013). Whatever the assessment method used, objective reliability stands as important requirement. Although proposals to control such variability was introduced earlier (Kroboth et al., 1992), a continued effort and search shall continue to ensure consistency and reproducibility of such process in the teaching and assessment of each discipline.

Oral case presentation is a vital component of teaching in the discipline of orthodontics. Due to the multiple elements required in its case presentations, the ambiguous level of knowledge display expected, and the increase number of students requiring multiple assessors, a form of rubric is needed to control the process of assessment. The primary aim of this study was to evaluate the reliability of using such a method in dental education, specifically for orthodontic oral case presentation for undergraduate students' assessment. As a secondary aim, potential correlations between instructors'/course director grading and the students' final course grade were investigated.

## 2. Materials and methods

During a series of three weeks orthodontic case presentation sessions, a new rubric-based case presentation assessment form (Fig. 1) was designed and introduced to three contributing instructors (Instructor A, B, and C). The form included three major categories concerning the quality of records, accuracy of data, and display of understanding the materials being presented. Each category was subdivided for two items for the ease of grading. A simple grading scale (grid) was displayed

**ORAL CASE PRESENTATION EVALUATION FORM**

Student's Name		Student's #	
<b>Category</b>	<b>Item</b>	<b>Student's Grade</b>	<b>Total Grade</b>
a. Quality of slides and records	Quality of pictures		<b>1</b>
	Quality of records		<b>1</b>
b. Accuracy of data presented	Correct clinical findings		<b>1</b>
	Correct Records Analysis		<b>1</b>
c. Understanding of data presented	Study model analysis		<b>1</b>
	Cephalometric analysis		<b>1</b>
<b>Total</b>			<b>6</b>
Instructor's Name:		Date:	

**Grading Rubrics:**

	<b>Satisfactory (1)</b>	<b>Compromised (7.5 - 0.5)</b>	<b>Poor (0.25 - 0.0)</b>
<b>Quality of pictures</b>	Complete set of good extra- & intra-oral pics representing the case	Missing some pics and/or compromised quality	Missing most pics and/or poor quality
<b>Quality of Records</b>	Complete views of study models & Ceph representing the case	Missing some views of study models/ Ceph	Missing most views of study models / Ceph
<b>Correct clinical findings</b>	Accurate Extra- & Intra-oral assessment's findings	Inaccurate some of the findings	Inaccurate most of the findings
<b>Correct Records Analysis</b>	Accurate study models & cephalometric analysis	Inaccurate some of the findings	Inaccurate most of the findings
<b>Study model analysis Comprehension</b>	Answering a question about study model analysis	Partially answering	Not answering
<b>Cephalometric analysis Comprehension</b>	Answering a question about ceph. analysis	Partially answering	Not answering

\*Examples of such questions include, but not limited, to:

- Your case shows a molar (canine) relationship in Class I (II or III), what if the mesiobuccal cusp (canine cusp) of upper molar drifted (mesially / distally).
- What is the interpretation of ANB angle of -6 degrees? What would be the interpretation if the upper incisor inclination to palatal plane was 130 degrees? ... etc.

A blank grading sheet has been supplied to the students for their guidance in preparation

**Fig. 1** Rubric-based orthodontic case presentation form used to evaluate students. Note categories distribution, grading guidelines, and the grading scale.

at the bottom of the page. Additionally, the form included grading guidelines that contained a sample of questions that can be asked during the presentation.

Prior to the beginning of the case presentation session, 5 min were spent with each instructor to introduce the form.

The forms were also presented and supplied to the studied fourth year male dental students early before they started working on their presentation preparation. In each instructor's group, the course director (C.D.), along with the assigned instructor, assessed the first eight students utilizing the same

assessment form (total of 24 students). The two final scorings made by the assigned instructor and the course director were then gathered for each student. The instructors were blinded to the fact that the course director was also taking part of the assessment process.

The data of this prospective comparative study was then analyzed using paired *t*-test to look for any significant differences in the scoring of the course director and each instructor in each group. Furthermore, Pearson's correlation was applied to test for any significant correlations between the instructors'/course director's grading and the students' final course grade. Also, correlations between the discrepancies of the instructors'/course director's grading (Instructor's grade – C. D. grade) and the students' final course grade was investigated. All statistical analysis was performed using SPSS program (Version 16 SPSS Inc., Chicago, IL), as part of internal course quality assurance initiative.

### 3. Results

Table 1 shows the mean grades scored for each pair of instructor and C.D. Some discrepancies were found, however, discrepancies were minimal and did not reflect any statistical significant when subjected to paired *t*-test analysis.

No significant correlations have been detected between the instructors'/course director's grading and the students' final course grade. Furthermore, correlations between the discrepancies of the instructors'/course director's grading (Instructor's grade – C.D. grade) and the students' final course grade did not show any significant correlations ( $p = 0.585$ ).

### 4. Discussion

The use of rubric is well documented in the literature for not only formal course assessment, but also in online course evaluation and development (Blood-Siegfried et al., 2008). The strength of this tool enabled its application in many aspects of the field of teaching. The present study attempted to investigate the reliability of its use to assess students' orthodontics oral case presentation. The use of rubric-based method was applied in assessing other students' case and research presentations (in pharmacy, and surgery fields respectively) and was found to be a successful tool (Musial et al., 2007; O'Brien et al., 2008).

The assessment form contained specific criteria listed in three categories. This is in order to serve as a guide for instructors' assessments and as a tool for student learning during their preparations. Therefore, a blank copy was introduced and sup-

plied to the students early before they started preparing their presentations. Generally, students appreciate clear guidelines to help them complete procedures. Repeated use of the rubric provides opportunities for student to achieve competency (Dong et al., 2011).

The results showed that there were no statistical significant differences between the scoring among each instructor and the course director. The form produced appeared to be reliable for assessing orthodontic case presentation. The process of grading was consistent, thus minimizing students' complaints from inconsistent grading. Kruger & Dunning (1999) demonstrated that the poor students tend to overestimate themselves, and vice versa. Using rubric would get students' attentions to appreciate their deficiencies from their own prospective (Kruger and Dunning, 1999). Thus minimizing discomfort and criticism that is associated with the feedback delivery process (Wood, 2000).

It was hypothesized that better performing students would display more consistent grading among different assessors. The present data however did not support this hypothesis. There was no significant correlation ( $p = 0.585$ ) between the instructors'/course director's grading discrepancies (Instructor's grade – C.D. grade) and the students' final course grade. Possibly due to the criteria listed in the form that controls the grading process regardless of the student's overall knowledge and course standing as reflected by his/her final course grade.

In order for a rubric to be more efficient, sometimes students should participate in the formulation of their own assessment rubric. This would result in a use of a language of both faculty members and students that would clearly represent expectations of each criterion and standard (Moni et al., 2005). Even that the results indicates reliable rubrics from the assessment side, a feedback must be always sought from student to aid in the upcoming revision process of the rubric. The presented rubric and its wording has been found to be fairly accepted by current students.

The mean grade the randomly selected students for the oral presentation ranged from 5.20 to 5.66 out of 6.00. This indicates that on average students have performed well in their presentations. This can be attributed to the rubric guidelines presented early for the students to set their expectations and guide them through the preparation process. The rubric supplied has served as instant student feedback tool. Providing effective feedback was shown to enable students to develop a deeper approach to their learning and improve learning outcomes (Mohanna and Chambers, 2003).

The present study illustrated the applicability and value of using rubric in orthodontic case presentation assessment. It would stimulate further studies to propose rubric methods to

**Table 1** Means of course director's and instructors' A, B, and C scorings along with p-value for paired *t*-test statistics being applied for each pair of assessors.

		Mean	N	S.D.	Std. Error Mean	p-Value
Pair 1	Course Director	5.656	8	0.229	0.081	0.129
	Instructor A	5.34	8	0.640	0.226	
Pair 2	Course Director	5.594	8	0.265	0.094	0.563
	Instructor B	5.656	8	0.265	0.094	
Pair 3	Course Director	5.388	8	0.179	0.063	0.208
	Instructor C	5.200	8	0.330	0.117	

be implemented in the teaching of orthodontics. Further studies can be initiated to re-evaluate the reliability of the presented form in different academic settings and larger sample size, given the limited current sample size number. Another limitation of the current study that can be considered in reproducing its method, is the inclusion of subjection from both genders. Furthermore, the current grading scale of the form (three categories), can be further expanded into four or five assessment's ranking.

## 5. Conclusion

The use of rubrics was found to be a promising reliable element to be included in the assessment of orthodontic courses. This study would stimulate the development of a valid well-defined orthodontic oral case presentation assessment form that can be tested on a larger group. Such form can be developed not only for instructor's use, but also for peer and self-assessment applications.

## References

- Baldwin, S.G., Harik, P., Keller, L.A., Clauser, B.E., Baldwin, P., Rebbecchi, T.A., 2009. Assessing the impact of modifications to the documentation component's scoring rubric and rater training on USMLE integrated clinical encounter scores. *Acad. Med.* 84 (10 Suppl), S97–100.
- Blommel, M.L., Abate, M.A., 2007. A rubric to assess critical literature evaluation skills. *Am. J. Pharm. Edu.* 71 (4), 63.
- Blood-Siegfried, J.E., Short, N.M., Rapp, C.G., Hill, E., Talbert, S., Skinner, J., Campbell, A., Goodwin, L., 2008. A rubric for improving the quality of online courses. *Int. J. Nurs. Edu. Scholar.* 5, 34.
- Bloom, B.S., Krathwohl, D.R., 1956. *Taxonomy of educational objectives: Handbook I, The cognitive domain.* David McKay & Co, New York.
- Curry, L., 1983. An organisation of learning styles theory and constructs. *Edu. Resour. Inform. Center* (235), 185.
- D'Antoni, A.V., Zipp, G.P., Olson, V.G., 2009. Interrater reliability of the mind map assessment rubric in a cohort of medical students. *BMC Med. Edu.* 9, 19.
- Daggett, L.M., 2008. A rubric for grading or editing student papers. *Nurse Edu.* 33 (2), 55–56.
- Dong, C., Asadoorian, J., Schönwetter, D.J., Lavigne, S.E., 2011. *Rubric Development Tools: Dentistry and Dental Hygiene Applications.* San Diego Faculty Development Workshops.
- Gadbury-Amyot, C.C., Kim, J., Palm, R.L., Mills, G.E., Noble, E., Overman, P.R., 2003. Validity and reliability of portfolio assessment of competency in a baccalaureate dental hygiene program. *J. Dent. Edu.* 67 (9), 991–1002.
- Guba, E.G., 1978. *Toward a methodology of naturalistic inquiry in educational evaluation.* Center for the Study of Evaluation, Los Angeles.
- Hancock, A.B., Brundage, S.B., 2010. Formative feedback, rubrics, and assessment of professional competency through a speech-language pathology graduate program. *J. Allied Health* 39 (2), 110–119.
- Kirkpatrick, D.L., 1979. *Techniques for evaluating training programs.* In: Ely, D.P., Plomp, T. (Eds.), *Classic Writings on Instructional Technology.* Libraries Unlimited, Inc., Englewood.
- Krathwohl, D.R., Bloom, B.S., Masia, B.B., 1964. *Taxonomy of Educational Objectives. The Classification of Educational Goals, Handbook II: Affective Domain.* David McKay Company, Inc., New York.
- Kroboth, F.J., Hanusa, B.H., Parker, S., Coulehan, J.L., Kapoor, W. N., Brown, F.H., Karpf, M., Levey, G.S., 1992. The inter-rater reliability and internal consistency of a clinical evaluation exercise. *J. Gen. Intern. Med.* 7 (2), 174–179.
- Kruger, J., Dunning, D., 1999. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessment. *J. Person & Soc. Psych.* 77 (6), 1121–1134.
- Lewin, L.O., Beraho, L., Dolan, S., Millstein, L., Bowman, D., 2013. Interrater reliability of an oral case presentation rating tool in a pediatric clerkship. *Teach. Learn. Med.* 25 (1), 31–38.
- Mohanna, K., Chambers, R., Wall, D., 2003. *Teaching Made Easier* Milton Keynes. Radcliffe Publishing Ltd.
- Moni, R.W., Beswick, E., Moni, K.B., 2005. Using student feedback to construct an assessment rubric for a concept map in physiology. *Adv. Physiol. Edu.* 29 (4), 197–203.
- Musial, J.L., Rubinfeld, I.S., Parker, A.O., Reickert, C.A., Adams, S. A., Rao, S., Shepard, A.D., 2007. Developing a scoring rubric for resident research presentations: a pilot study. *J. Surg. Res.* 142 (2), 304–307.
- O'Brien, C.E., Franks, A.M., Stowe, C.D., 2008. Multiple rubric-based assessments of student case presentations. *Am. J. Pharm. Edu.* 72 (3), 58.
- Pangaro, L.A., 1999. New vocabulary and other innovations for improving descriptive in-training evaluations. *Acad. Med.* 74, 1203–1207.
- Ramsden, P., 1992. *Learning to Teach in Higher Education.* Routledge, London.
- Scriven, M., 1991. *Evaluation Thesaurus.* Sage, Newbury Park, CA.
- Shapiro, I., 2001. Doctor means teacher. *Acad. Med.* 76 (7), 711.
- Simpson, E.J., 1972. The classification of educational objectives in the psychomotor domain. *The Psychomotor Domain.* 3:43-56. Silver Spring, Gryphon House, Inc.
- Spencer, J., 2003. Learning and teaching in the clinical environment. *BMJ* 326 (7389), 591–594.
- Speyer, R., Pils, W., Van Der Kruis, J., Brunings, J.W., 2011. Reliability and validity of student peer assessment in medical education: a systematic review. *Med. Teach.* 33 (11), e572–585.
- Stufflebeam, D.L., 2003. The CIPP Model for Evaluation, in: Kellaghan, T., Stufflebeam, D.L. (Eds.), *International Handbook of Educational Evaluation.*
- Turbow, D.J., Werner, T.P., Lowe, E., Vu, H.Q., 2016. Norming a written communication rubric in a graduate health science course. *J. Allied Health* 45 (3), e37–42.
- Tyler, R.W., 1949. *Basic Principles of Curriculum and Instruction.* The University of Chicago Press, Chicago.
- White, J.G., Kruger, C., Snyman, W.D., 2008. Development and implementation of communication skills in dentistry: an example from South Africa. *Eur. J. Dent Edu.* 12 (1), 29–34.
- Wiese, J., Varosy, P., Tierney, L., 2002. Improving oral presentation skills with a clinical reasoning curriculum: a prospective controlled study. *Am. J. Med.* 112 (3), 212–218.
- Wood, B.P., 2000. Feedback: a key feature of medical training. *Radiology* 215 (1), 17–19.