

A New Holistic Assessment System and its Impacts on Student Performance in Regional Anatomy

Un Nuevo Sistema de Evaluación Holística y sus Impactos en el Rendimiento de los Estudiantes en Anatomía Regional

Abudureyimujiang Ruze; Siyiti Amuti; Niu Lipan & Fengxia Liu

RUZE, A.; AMUTI, S.; LIPAN, N. & LIU, F. A new holistic assessment system and its impacts on student performance in regional anatomy. *Int. J. Morphol.*, 38(4):863-868, 2020.

SUMMARY: Summative and formative assessments are two overlapping complementary ways of assessing student progress in regional anatomy teaching and learning. Our present study was designed to create a new holistic assessment system that embraces both summative and formative assessments, and evaluate its impacts on student performance of regional anatomy. A collection of five formative assessment tests were designed and introduced into the teaching process of regional anatomy, and their performances were combined with scores of the summative assessment taken at the end of the semester to form the holistic assessment. And an anonymous survey was conducted to gather student perceptions regarding the assessments. We found, compared to summative assessment scores, students' overall average points are higher by 4.67 points ($P < 0.05$) and 9.23 points ($P < 0.01$) when evaluated by the holistic and formative assessment; formative assessment scores are positively correlated to summative assessment scores, and the Pearson correlation is 0.624. Questionnaire investigation showed 57.65 % of the students wishes to be assessed by the holistic assessment, and 97.9 % of the students think that the holistic assessment can promote the frequency of student-teacher communication, which helps them form the right learning attitude and improve the performance. The results indicated that holistic assessment is a more reflective and practical approach of evaluating student performance in regional anatomy teaching, which can increase student-teacher communication and enhance the self-directed learning among students.

KEY WORDS: Regional anatomy; Formative assessment; Summative assessment; Holistic assessment.

INTRODUCTION

Regional anatomy is a keystone, both in teaching medical curricula and in clinical practice, it provides the building blocks that underpin the foundations for understanding many pathological processes and their management. A reduction in the number of hours dedicated to regional anatomy teaching urged anatomists to explore and identify time efficient techniques for enhancing the teaching efficiency and the students' performance (Caswell *et al.*, 2015). Teaching is a complex process that needs students and teachers to both participate in and accomplish together (Joyce *et al.*, 2018). Competency-based education requires acknowledging that teaching and learning are not synonymous, teachers are tasked with ensuring an optimal learning environment and measuring student progress toward achieving learning outcomes (DiVall *et al.*, 2014). The single most desirable

improvement in regional anatomy teaching would probably be in the field of evaluation. During teaching process, teachers need to implement various types of assessments to evaluate, and enhance both teaching and learning efficiency.

Summative assessment (SA) and formative assessment (FA) are two overlapping complementary ways of assessing student progress. SA aims to evaluate student learning and academic achievement at the end of a year or semester by comparing it against a universal standard or school benchmark, it not only tests the final effect of student learning, but more importantly, evaluates the learning consciousness of students and their ability to integrate the whole course. But when feedback is provided in a SA context, it is not always used effectively by

learners. FA refers to an approach that evaluates student performance timely, dynamically and repeatedly (DiVall *et al.*; Evans *et al.*, 2014; Del Fiol *et al.*, 2016). It consists of tests and activities to determine the level of knowledge of students for the purpose of providing feedback and planning of future instruction of self-directed learning, then students know how and what to learn (Caswell *et al.*; Gerhard-Szep *et al.*, 2016). FA is more diagnostic than evaluative, usually systematic in approach, and is designed to be available to students during a particular period of study to provide motivation for learning (Mitra & Barua, 2015). It also familiarizes students with the expectations of SA, required levels of learning and provides feedback that guides the direction of both the student learning and a faculty member's teaching (Kibble, 2017).

However, we found that students tend to stop reviewing the related knowledges after periodic exams, and they lose the willingness and the ability to integrate various parts of the course, and lack of systematic understanding of the course. Therefore, in order to achieve high efficiency in teaching and learning of regional anatomy, we introduced a series of FA tests into the teaching process of regional anatomy, through which timely feedback is given to students so that they can adjust their learning strategies accordingly in each module. FA not only fulfills the needs of the learner, but it also fulfills a teacher's obligations to students by providing appropriate resources to aid in learning and teaching. But in most cases, FA often has low or no point value, which means it is still the SA that provides a final measure of student performance, and it is still the most conventional and generally accepted way of evaluating student-learning in China and many other countries in the world. However, the fact that student performance in FA doesn't contribute to the final measure not only undermines its value, and but may also decrease student participation and the time students spend on respective subject. Knowing that FA tests does not contribute to the final measuring system, students might consciously or unconsciously tend to decrease the importance they attach to FA tests.

We believe that a holistic assessment (HA) consisted of both FA and SA combined appropriately might be a better measuring system in terms of being practical and reflective of student performance than any other assessment systems, FA or SA, implemented alone in teaching and learning of regional anatomy. And this study was conducted to compare three different types of evaluation systems (HA, FA, and SA) and explore their effect on student performance of regional anatomy.

MATERIAL AND METHOD

The institutional review board of Xinjiang Medical University granted this educational method study an exemption from formal review (Educational Teaching Reform and Research Foundation of Xinjiang Medical University, YG2017053). Permission was obtained from all students to use de-identified performance assessments for this study and subsequent publications. A total of 426 third year students participated in the study across 8 separate iterations of tests that contain 4 module tests (4 quizzes, limbs, head and neck, abdomen, thorax and pelvis), 2 spotter tests using specimens (mid-term and final specimen exams), a mid-term exam and a final summative exam (September 2016–January 2017).

Establishment of the FA. A collection of five FA test types including quizzes during lectures, after-school assignments, four module tests, two spotter tests using specimens and a mid-term exam were designed and introduced into the teaching process. Among them, the former two (quizzes during lectures and after-school assignments) does not contribute to the overall average FA scores, and the rest of the FA tests contribute evenly to the overall average FA scores. At the end of the course, student performane were evaluated sololy by the overall average FA scores. The FA tests include Among them, the former two (quizzes during lectures and after-school assignments) does not contribute to the HA, they were purely designed to diagnose and monitor how the students were responding to the lectures and lab class being presented. And the latter three (four module tests, two spotter tests and a mid-term exam) account for 10 % for each in HA. The SA test refers to the summative exam taken at the end of the semester to test students' overall grasp and understanding of the whole knowledge of the book and makes up 70 % of the HA.

Establishment of the SA. The SA test is formed merely by one single exam taken at the end of the course to test students' overall grasp and understanding of the whole knowledge of the book. At the end of the course, student performance were evaluated purely by the exam taken at the end of the course.

Establishment of the HA: HA is a synthetic complex of FA and SA tests. It is formed by the SA and FA scores meshed up at a certain proportion. As it is mentioned above, FA tests include quizzes during lectures, after-school assignments, four module tests, two spotter tests using specimens and a mid-term exam. Among them, the former two (quizzes during lectures and after-school assignments) does not contribute to the HA, they were purely designed

to diagnose and monitor how the students were responding to the lectures and lab class being presented. The latter three (four module tests, two spotter tests and a mid-term exam) account for 10 % for each in HA. The SA test refers to the summative exam taken at the end of the semester to test students' overall grasp and understanding of the whole knowledge of the book and it makes up 70 % of the HA. Table I provides a detailed summary of the FA tests and SA that released, and illustrates their contents, time of release, their proportion in the HA, and details of when students receive answers and feedback.

Questionnaire: Student opinions were gathered from a handout on-site paper-based survey asking 6 questions

about their experience of the measuring systems used during the teaching process. 407 students out of 426 responded to the survey. In the first question we asked students "By which assessment system do you wish to be evaluated?" to find out which type of evaluation system is more appealing to students. And under the hypothesis that the reason FA benefits student learning is that it mainly enhances student-teacher communication, we asked students about the ideal and actual rate of student-teacher communication in question number two and three. To explore to what aspect the feedback from FA contribute to, we asked students about impact of the feedback from FA in the last two question. The questions asked and their respective answer choices in the survey are included in Table II.

Table I. The contents of regional anatomy teaching evaluation system.

Assessment	Proportion	Contents	Answer/feedback
Quizzes during lectures	0.00	Interactive questions asked during lectures, each student have at least one chance to participate. It was used in each lecture in order to break up the session, to see how they were responding to the material being presented.	Immediate
After-school assignments	0.00	Essay questions about the key points of the lecture and a word list with which the students should become familiar with and be able to define were assigned.	One week after release
Module tests	0.10	Four separate tests regarding head&neck, limbs, chest&abdomen, pelvic&perineum were taken to test students' grasp of knowledge of respective parts.	One week after release
Spotter tests using specimens	0.10	Two spotter tests at mid-term and at the end of the semester using specimens were taken to test students' ability to use key knowledge to identify related parts and structures on human specimens.	Immediate
Mid-term exam	0.10	Taken in the mid-part of the semester to test students' knowledge and understanding of the first half of the book.	One week after release
Summative exam	0.70	Taken at the end of the semester to test students' overall grasp and understanding of the whole knowledge of the book.	One week after release

Table II. The questions asked in the survey.

Questions used in survey	Answer choices(%)			
	FA	SA	HA	Other systems
By which assessment system do you wish to be evaluated?	13.57 %	26.89	57.65 %	1.89 %
The amount of student-teacher communication should be	Excessive 34.10 %	Relatively a lot 43.52 %	Normal amount 20.25 %	Little 2.1 %
How often do you engage in an active communication with your teacher?	Very often 12.79 %	Occasionally 44.22 %	Normally not 39.43 %	Never 3.55 %
Formative assessment enhances student-teacher communication?	Strongly agree 32.78 %	Agree 57.21 %	Disagree 8.81 %	Strongly disagree 1.20 %
With which areas the does the feedback from formative assessment help?	Learning methods 32.14 %	Learning attitude 45.29 %	Doesn't help 21.31 %	Other areas 1.24 %
Formative assessment can contribute my learning efficiency	Strongly agree 14.20 %	Agree 53.46 %	Disagree 22.02 %	Strongly disagree 10.30 %

Data analysis: The data collected were tabulated and analyzed by using the Statistical Package for Social Sciences (SPSS) version 22.0 for windows. The comparison between the average scores of SA, FA and HA was conducted by the statistical test of one way anova. And a Pearson correlation was conducted to analyze the correlation FA and SA scores, between SA and HA scores, FA and HA scores. In this study?statistical test assumptions were verified and P values less than 0.05 were considered significant.

RESULTS

Students' average scores under three different assessment systems: Students' average scores under three different assessment systems (FA, SA and HA) were illustrated in Table III. Compared to SA, students' average points increased by 9.23 points ($P < 0.01$) and 4.47 points ($P < 0.05$) when evaluated by the FA and HA. The results showed the students' performance is higher when evaluated by the FA, but lower when evaluated by the SA, and moderately by the HA.

Correlation between three assessment systems: FA scores positively correlates to SA scores, and the Pearson correlation is 0.624 ($P < 0.05$, Table IV). This means that those who do well in their FA tests might also get a higher grade in their SA tests and those who fail on their FA tests might also likely to fail on the summative test. SA scores positively correlates to HA scores, and the Pearson correlation is 0.998 ($P < 0.05$, Table IV).

Students' perceptions of HA and FA: A total of 407 out of 426 students responded to the on-site survey, giving a response rate of 95.08 %. The results are summarized in Table II. 57.65 % of the students wishes to be assessed by the HA. 97.9 % of the students think that the frequency of student-teacher communication should be at a higher or at least at a normal rate, while only 12.79 % of the students engage in active communications with teachers on a frequent basis and 44.22 % on occasional basis. On the other hand, others

Table III. Students' average scores under three different assessment systems.

	FA	SA	HA
Average scores	75.57 ±13.35	66.34 ±16.79 ^F	71.01 ±14.12 ^{FS}

Means $P < 0.01$ when compared to SA; Means $P < 0.01$ when compared to FA.

Table IV. Pearson correlation between FA and SA.

Group		FA	SA
FA	Pearson Correlation	1	0.624
	P		0.0001
	n	370	370

normally don't or never communicate with teachers. After introducing FA, 89.99 % of the students agrees or totally agrees with the fact that FA enhances student-teacher communication. 32.14 % of the students think that FA helps them find better learning methods and 45.29 % of the students think that FA helps them form the right learning attitude.

DISCUSSION

Students' affinity for acquiring and retaining knowledge is likely determined by alignment of intrinsic learning styles and surrounding learning environments (Chapman & Hakeem, 2015). Progress assessments, both FA and SA, are deemed imperative part of the educational process.

The goal of SA is to evaluate student learning at the end of a school year or semester by comparing it against certain standard or benchmark (Fowell *et al.*, 1999; Raupach *et al.*, 2013; Mitra & Barua). Since SA focuses mainly on the single test or output at the end, it not only makes almost all individuals anxious and disruptive, but also might lead students to spend great deal of time and energy on test papers that resembles the SA instead of focusing on the developing practical skills in regional anatomy. Besides, being a single test, there is no chance for the students to recover or compensate (Schaper *et al.*, 2013).

FA, on the other hand, refers to a wide variety of methods that teachers use to conduct in-process evaluations of student comprehension, learning needs, and academic progress during a lesson or unit (Waghlikar *et al.*, 2013). FA help teachers identify concepts that students are struggling to understand, skills they are having difficulty acquiring, or learning standards they have not yet achieved so that adjustments can be made to lessons, instructional techniques, and academic support.

FA deals with small areas of content as it is an ongoing process (Lee *et al.*, 2018). While, SA deals with the whole project as it is performed after the completion of a certain course. SA can be used to great effect in conjunction and alignment with FA, and instructors can consider a variety of ways to combine these approaches. The primary purpose of this study was to determine if HA, the combination of SA and FA, is a better measuring system in terms of being practical and being reflective of student performance than SA or FA implemented alone. Researchers have found that FA would enhance summative exam scores in medical students, dental students, and variety of other undergraduate majors (Brar *et al.*, 2007; DiVall *et al.*; Chisnall *et al.*, 2015; Valero & Cardenas, 2017).

Our study illustrates that students' average scores increases when the FA accounts for a certain portion in the final measuring system. And when the FA was solely used as the final measuring system, students' average scores were the highest. This outcome does not necessarily mean that FA, when used as the only final measure would increase student performance or is the most reflective form of measurement. Because FA tests in our study were consisted of several separate modules taken in different period of times during the semester, they only deals with the small areas of content as it an ongoing process. They can only evaluate student performance in respective module. Therefore, FA tests can't measure the overall grasp or the understanding of the whole knowledge of the book. One of the possible reasons for this phenomenon is the fact that FA tests require a much smaller range of knowledge than the SA test. And this makes it easier for students to prepare for FA test. Therefore, when FA tests accounted for the 30 % of the final measuring system, students' average scores was increased by 4.67 points ($P < 0.05$). And when FA was used as the only final measuring system, students' average scores was increased by 9.23 points ($P < 0.01$). This study shows that students' FA scores positively correlate to their SA scores, meaning FA scores, in some extent, can predict SA scores. Reasons for the improvement of performance between FA and SA may include increased motivation from a poor mark in FA tests, improved exam techniques from the chances to practice and the findings of better learning strategies. Our study also indicates that SA scores positively correlate to HA scores. This means that, being a major element in HA, SA enables HA to reach the goal of evaluating student learning at the end of a school year or semester by comparing it against certain standard or benchmark.

Questionnaire from this study reveals that, most of the students expects a higher rate of student-teacher communication than they actually engage in and the implementation of FA increased the frequency of it. This may be one of the reasons FA would enhance SA scores. On the other hand, there are also reports that argue, in fact, FA does not improve summative exam performance (Buchanan, 2000; Brothen & Wambach, 2001; Haberyan, 2003). One of the underlying reasons for that is the fact that student performance from FA usually does not contribute to the overall measuring system undermines its value and lessens the importance students give to it (Fowell *et al.*; Cook *et al.*, 2006; Chisnall *et al.*). This mechanism would result in a situation that students may go through the motions and not give their best when taking FA tests, meaning students who pass the FA tests are likely to pass the SA, but the majority of students who fail the FA tests are also likely to pass the SA. Therefore adding the test scores from FA to the overall measuring system would encourage students to attach greater

importance to FA tests, which in turn promotes student learning.

In this study, the final measuring system, HA is a hybrid of SA and FA, in which SA and FA respectively accounts for the 70 % and 30 %. As it has been listed on the Table I, FA also includes quizzes during lectures and after-school assignments, which accounts for 0 % of the final measurement. In spite of their zero percentage, timely feedback is given to the students. As for the proportions of the assessments in HA, SA accounts for 70 % because the final measuring system should be able to evaluate students' overall grasp and understanding of the whole knowledge of the subject. Three sets of FA tests account for 30 %, because adding the test scores from FA to the final measuring system would encourage students to attach greater importance to FA tests, which in turn promotes student learning.

ACKNOWLEDGEMENTS

The authors are grateful to the Research and Ethics Committee of Xinjiang Medical University for granting permission to conduct this study.

RUZE, A.; AMUTI, S.; LIPAN, N. & LIU, F. Un nuevo sistema de evaluación holística y sus impactos en el rendimiento de los estudiantes en anatomía regional. *Int. J. Morphol.*, 38(4):863-868, 2020.

RESUMEN: Las evaluaciones sumativas y formativas son dos formas complementarias superpuestas de evaluar el progreso de los estudiantes en la enseñanza y el aprendizaje de la anatomía regional. El presente estudio fue diseñado para crear un sistema nuevo de evaluación integral que abarque tanto las evaluaciones sumativas como las formativas, y evalúe sus impactos en el rendimiento de los estudiantes de la anatomía regional. Se diseñó e introdujo una colección de cinco pruebas de evaluación formativa en el proceso de enseñanza de la anatomía regional, y sus desempeños se combinaron con los puntajes de la evaluación sumativa tomada al final del semestre para formar la evaluación holística. Además, se realizó una encuesta anónima para recopilar las percepciones de los estudiantes con respecto a las evaluaciones. Encontramos que, en comparación con los puntajes de la evaluación sumativa, los puntos promedio generales de los estudiantes son más altos en 4,67 puntos ($P < 0,05$) y 9,23 puntos ($P < 0,01$) cuando se evalúan mediante la evaluación holística y formativa; los puntajes de las evaluaciones formativas se correlacionan positivamente con los puntajes de las evaluaciones sumativas, y la correlación de Pearson es 0,624. La investigación del cuestionario mostró que el 57,65 % de los estudiantes desea ser evaluado por los evaluadores holísticos, y el 97,9 % de los estudiantes piensa que la evaluación holística puede promover la frecuencia de la comunicación entre estudiantes y maestros, útil para formar una actitud correcta de

aprendizaje y mejorar el rendimiento. Los resultados indicaron que la evaluación holística es un enfoque más reflexivo y práctico para evaluar el desempeño de los estudiantes en la enseñanza de anatomía regional, lo que puede aumentar la comunicación entre estudiantes y maestros y mejorar el aprendizaje autodirigido entre los estudiantes.

PALABRAS CLAVE: Anatomía regional; Evaluación formativa; Evaluación sumativa; Evaluación holística.

REFERENCES

- Brar, M. K.; Laube, D. W. & Bett, G. C. L. Effect of quantitative feedback on student performance on the National Board Medical Examination in an obstetrics and gynecology clerkship. *Am. J. Obstet. Gynecol.*, 197(5):530.e1-5, 2007.
- Brothen, T. & Wambach, C. Effective student use of computerized quizzes. *Teach. Psychol.*, 28(4):292-4, 2001.
- Buchanan, T. The efficacy of a World-Wide Web mediated formative assessment. *J. Comput. Assist. Learn.*, 16:193-200, 2000.
- Caswell, F. R.; Venkatesh, A. & Denison, A. R. Twelve tips for enhancing anatomy teaching and learning using radiology. *Med. Teach.*, 37(12):1067-71, 2015.
- Chapman, S. J. & Hakeem, A. R. Undergraduate anatomy teaching: evaluations and conclusions. *Acad. Med.*, 90(4):399, 2015.
- Chisnall, B.; Vince, T.; Hall, S. & Tribe, R. Evaluation of outcomes of a formative objective structured clinical examination for second-year UK medical students. *Int. J. Med. Educ.*, 6:76-83, 2015.
- Cook, D. A.; Thompson, W. G.; Thomas, K. G.; Thomas, M. R. & Pankratz, V. S. Impact of self-assessment questions and learning styles in Web-based learning: a randomized, controlled, crossover trial. *Acad. Med.*, 81(3):231-8, 2006.
- Del Fiol, G.; Mostafa, J.; Pu, D.; Medlin, R.; Slager, S.; Jonnalagadda, S. R. & Weir, C. R. Formative evaluation of a patient-specific clinical knowledge summarization tool. *Int. J. Med. Inform.*, 86:126-34, 2016.
- DiVall, M. V.; Alston, G. L.; Bird, E.; Buring, S. M.; Kelley, K. A.; Murphy, N. L.; Schlesselman, L. S.; Stowe, C. D. & Szilagyi, J. E. A faculty toolkit for formative assessment in pharmacy education. *Am. J. Pharm. Educ.*, 78(9):160, 2014.
- Evans, D. J.; Zeun, P. & Stanier, R. A. Motivating student learning using a formative assessment journey. *J. Anat.*, 224(3):296-303, 2014.
- Fowell, S. L.; Southgate, L. J. & Bligh, J. G. Evaluating assessment: the missing link? *Med. Educ.*, 33(4):276-81, 1999.
- Gerhard-Szep, S.; Güntsch, A.; Pospiech, P.; Söhnel, A.; Scheutzel, P.; Wassmann, T. & Zahn, T. Assessment formats in dental medicine: An overview. *GMS J. Med. Educ.*, 33(4):Doc65, 2016.
- Haberyan, K. A. Do weekly quizzes improve student performance on general biology exams? *Am. Biol. Teach.*, 65(2):110-4, 2003.
- Joyce, J.; Gitomer, D. H. & Iaconangelo, C. J. Classroom assignments as measures of teaching quality. *Learn. Instr.*, 54:48-61, 2018.
- Kibble, J. D. Best practices in summative assessment. *Adv. Physiol. Educ.*, 41(1):110-9, 2017.
- Lee, A.; Sandvei, M.; Asmussen, H. C.; Skougaard, M.; Macdonald, J.; Zavada, J.; Bliddal, H.; Taylor, P. C. & Gudbergesen, H. The development of complex digital health solutions: formative evaluation combining different methodologies. *JMIR Res. Protoc.*, 7(7):e165, 2018.
- Mitra, N. K. & Barua, A. Effect of online formative assessment on summative performance in integrated musculoskeletal system module. *BMC Med. Educ.*, 15:29, 2015.
- Raupach, T.; Brown, J.; Anders, S.; Hasenfuss, G. & Harendza, S. Summative assessments are more powerful drivers of student learning than resource intensive teaching formats. *BMC Med.*, 11:61, 2013.
- Schaper, E.; Tipold, A. & Ehlers, J. P. Use of key feature questions in summative assessment of veterinary medicine students. *Ir. Vet. J.*, 66(1):3, 2013.
- Valero, G. & Cárdenas, P. Formative and Summative Assessment in Veterinary Pathology and Other Courses at a Mexican Veterinary College. *J. Vet. Med. Educ.*, 44(2):331-7, 2017.
- Waghlikar, K. B.; MacLaughlin, K. L.; Kastner, T. M.; Casey, P. M.; Henry, M.; Greenes, R. A.; Liu, H. & Chaudhry, R. Formative evaluation of the accuracy of a clinical decision support system for cervical cancer screening. *J. Am. Med. Inform. Assoc.*, 20(4):749-57, 2013.

Corresponding author:

Dr. Fengxia Liu
Department of Human Anatomy
School of Basic Medical Science
Xinjiang Medical University
393 Xinyi Road
Urumqi 830011
Xinjiang Province
CHINA

E-mail: liufengxia555@126.com

Received: 30-10-2019

Accepted: 30-01-2020