Original article:
"Innovative Anatomy assessment methods in COVID-19 Pandemic: statistical observations and students viewpoints"

Nidhi Gupta[1], Suniti Pandey[2], Ashutosh Anshu[3]

Abstract:

Background: Currently anatomy training and assessment is seriously disrupted in COVID-19 Pandemic lockdown. New challenges of anatomy curriculum and assessment method development are being attempted based upon resources, expertise and technical skills of educators. This study aims at creating a standard operation procedure (SOP) for conducting gross anatomy assessment, finding statistical relevance and discusses viewpoints of students reflected in feedback. **Materials and methods:** a retrospective cross sectional study was conducted on 250 first year students at GSVM Medical College Kanpur India. Students were briefed about SOP and were given choice to opt for either of three interview modalities of assessment tool [Viz. Telephonic Interview (GROUP T), Whatsapp© video call (GROUP W) and Zoom© meetings (GROUP Z)] and submit subjective feedbacks to assessor regarding new assessment format. The marks obtained as response distribution were subjected to descriptive statistics. One way ANOVA was applied to test any difference in means at (P<0.05) followed by post-hoc pair wise test amongst groups corrected to *Bonferroni adjusted alpha level* of [P<0.0167] per test. **Results:** Among 97.2% respondent ANOVA test result pointed that the types of interview pattern preferred had no significant impact on obtained score [F (2,240) = 2.80, p = .062], post-hoc test further confirmed that there was no significant difference between groups. **Conclusion:** the assessor and students both need to evolve and adapt to paradigm changes. Student's feedback will serve to tailor the assessment tool by refining validity, consistency, reliability and interaction dynamics of assessment designs. Alternative evidence based futuristic assessment methods exploration are call of the hour.

Keywords: ANATOMY ASSESSMENT; ANOVA ; TELEPHONE; WHATSAPP© ; ZOOM©

Introduction

Formal anatomy training is the backbone of medical education.[1] During academic trainings, the thought process and information content of the students needs continuous and intermittent assessment. Presently there is drastic change in world owing to COVID-19 pandemic. The entire curriculum follow-up and academic training is seriously disrupted.[2] New challenges of effective transfer of course content through online platforms are delivered. To assess them precisely for academic records in this scenario is even harder.[2] The anatomy educators themselves are at helm of challenges and struggling with highly unfamiliar situation. More so, gross anatomy assessment is cumbersome as compared to simpler histology, neuro-anatomy, embryology and osteology, which are traditionally assessed through identifying and explaining slides, pointers and models.[3,4] To add to perplexity of situations, problem based learning (PBL) has to be incorporated

1. Dr. Nidhi Gupta MBBS, MS (ANATOMY), Assistant Professor, Department of Anatomy, GSVM Medical College, Kanpur 208002. INDIA
2. Dr. Suniti Pandey MBBS, MS (ANATOMY) , Professor and Department Head, Department of Anatomy, GSVM Medical College, Kanpur 208002. INDIA
3. Dr. Ashutosh Anshu MBBS, MS (ANATOMY), D.A.,Consultant Anaesthetist, Anatomist & Intensivist Ex Alumni & Visiting Faculty Department of Anatomy GSVM Medical College, Kanpur 208002. INDIA

Correspondence to: Dr. Nidhi Gupta MBBS, MS (ANATOMY), Assistant Professor, Department of Anatomy, GSVM Medical College, Kanpur 208002. INDIA. E-mail : drnidhianat@gmail.com

in teaching and assessment styles.[5] Students these days are devoid of face to face interaction vide administrative orders of lock-down. A viable option for the educator is assess gross anatomy through prosected specimens and short PBL questions through interactive telephonic or video call.[6-8] A significant amount of students do not have robust internet access, high end smart phones and laptops besides being sufferer of problems of accommodation and fooding. Procrastination and reluctance to active learning has swept in current gloomy environment. Educators are under pressure of tight declining schedules and timely curriculum as a lot remain to done in current academic calendar. The entire erstwhile framework of study and assessment pattern is relentlessly compromised. Hence, educators need to devise a tailored approach for assessment of students. Unlike being directive and didactic previously, teachers must now accommodate students view on methods of assessment.

Aim of the study:

We aimed at deriving inferences of what is needed to be done at creating an assessment model based upon feasibility to conduct tests as per inputs provided by students as a standard operation procedure (SOP). Therefore, this study focuses on statistically comparing three different assessment methods [Viz. Telephonic Interview, Whatsapp video call (Whatsapp© Facebook Inc., Mountain View, CA) and Zoom meetings (Zoom Video Communications, Inc©. San Jose, CA)], with nearly identical questionnaire for interviewing. The goal is to analyze the effect of each mode, including pinpointing differences in the attitude profile of students. We need to see what can deliver representative results. The outcome of this study will help in making mode choices and in interpretation of results of assessment methods. So far we know, such comparisons associated with the quality and representativeness of academic human anatomy assessment modes is scarce in literature.

Material and methods:

After continual discussion for maintaining educational in background of current disruptions, peer faculties in the department decided to continue on with assessment to keep pace with rapidly depleting curriculum time in academic calendar. A consensus was drawn for methodology of assessment. The cardinal features of SOP for assessment were:-

- The assessment of gross anatomy dissection of recently taught abdominal region was conducted through modified dissection part completion test (DPCT) as used to be in the past.

- Each Anatomy Faculty in the department procured certain number of questions pertaining to examine the students from thus creating common pool of questionnaire. Random sets of questions from this questionnaire pool were mixed and grouped into six separate bundles, one for each examiner. The confidentiality of prepared question bundles was secured and was given to assessor only during assessment time.

- 10 questions were asked for 2marks each for a total of 20 marks.

- The questions were classified into:-

➢ 2 straightforward questions for identification of structures in pre dissected specimens for those availing video calls/and for those attending telephonic call needed to define relations of structure as memory based factual questions.

➢ 2 questions about vascular/ nerve supply of viscera/structure.

➢ 2 questions about applied importance of a structure.

➢ 2 questions about diagnosing the involvement of a structure in particular set of situations, as PBL.

➢ 1 question for naming/identifying of one congenital anomaly of a particular viscera/ structure as PBL.

➢ 1 question for defining surface anatomy of a particular viscera/structure as PBL based on factual memory.

- The outline of assessment and scoring method was conveyed through a demonstration video created by the examiner to class representatives well in advance of assessment date. It was further ensured that it was circulated amongst every peer faculty and students.

- Date of assessment was announced 4 days prior to test after verification through class representative that each student has received information about briefings and priming for assessment. Students were advised to seek quiet place and focus at interview schedule. They were also conveyed to be ready for producing brief answer in time frame and not be explanatory unnecessarily. It was suggested to take a margin of time before and after interview as assessor will reach them.

- The duration of interactive session was fixed, and students were primed for allocating time for assessment.

- Now the students were left freely to choose interaction session through telephone calls, Whatsapp video call and Zoom meetings digital platform.[9-11] The students were supposed to brief

their choice for modalities of assessment through class representative with 2 positive or negative reasons at least 2 days prior to scheduled assessment days.

- Through class representative contact numbers of all students was made available to the representative faculty. Subsequently, the list containing contact and choice of session was circulated amid faculties.
- Those favoring Zoom interactive platform, were given meeting schedule invitation ID beforehand.

One day prior to assessment session every steps were cross checked.Faculties reached department beforehand and arranged specimens and necessary audiovisual settings. The administrative orders of lockdown and safe social distancing were followed throughout the session.

The interviews were all conducted on the same day during a decided allocated time slot. Time limits for response inputs were fixed through manual buzzer by examiner. The assessment was limited to 15 min of interaction of faculty per student regarding relevant topics.

The marks obtained by the students against roll number were recorded through pen/paper and the modality of assessment was written clearly by the assessor. Subsequently after completion of assessment, the marks obtained by three groups were constructed in the spreadsheet. Later those records were confidentially handed over to statistician for further analysis.

In order to test the effect of assessment modes in students, the inputs were categorized in three groups namely:-

- Students interviewed via Telephonic call (GROUP T)
- Students interviewed via Whatsapp Video call (GROUP W)
- Students interviewed via Zoom meeting platform (GROUP Z).

This study analyses statistical implications of test scores of students taking interview over either of three options. Response distribution was assessed by descriptive statistical analysis. *One-way analysis of variance ANOVA* was employed to test mean test scores difference between all groups, to accept/reject H1 (null: no overall mean test score differences exist between three groups). The threshold for statistical significance was set at ($P<0.05$). The obtained *ANOVA* test was subjected to *post-hoc test* for pair wise comparison of every group to find any statistical significant difference between mean test scores for

each groups in relation to one another, to accept/ reject H2(null: no pair wise significant differences exists between mean test scores for each groups). We conducted *two sample T-Test* assuming equal variance for this purpose. To control for family wise error rate across post-hoc comparisons (3 tests in total), a *Bonferroni adjusted alpha level* of [$P<0.05/3 =P<0.0167$] per test was employed as a threshold for significance.[12] Box and whisker plots were constructed to evaluate the overall trends of performance. Statistical analyses were carried out using IBM SPSS (v. 20; IBM SPSS, Armonk, NY).

This retrospective cross sectional study was done on 250 students of First year in Anatomy Department of GSVM Medical College Kanpur after administrative orders of country wide lockdown owing to Pandemic. Voluntary non random sampling technique was used for the purpose owing to limitations of resources available with the students.[13]

RESULTS:

In total, 243 (out of a total eligible population of 250 students) completed the assessment, indicating a response rate of 97.2%. Respondents were least in GROUP T 96.74% (119/123), moderate in GROUP W 97.53 % (79/81) and maximum in GROUP Z 97.82 (45/46). Two students in GROUP T did not receive the call while other two had multiple disconnections and poor voice access during interview. Two students in GROUP W and one student in GROUP Z did not receive Whatsapp/ Zoom meeting call. These seven students were marked absent in the conducted interview.

The descriptive statistics associated with students test scores in three different groups are reported in Table 1. It can be seen that GROUP T had numerically smallest mean level with maximum SD (12.20±1.92) and GROUP W had numerically highest mean level with least SD (12.79±1.45).

| Table 1 Descriptive statistics for assessment scores in three different groups			
	GROUP T	GROUP W	GROUP Z
No. of Observation	119	79	45
Mean	12.20	12.79	12.37
Median	12	13	12
Standard Deviation (SD)	±1.92	±1.45	±1.68
Standard Error	0.17	0.16	0.25

Table 1 Descriptive statistics for assessment scores in three different groups			
Kurtosis			
	0.45	0.27	0.06
Skewness			
	0.50	0.41	0.38
Standard error of Mean	0.51	0.4	0.53
Range	10	7	8
Confidence Level at (95.0%)	0.347	0.325	0.505
UL CI (95%)	12.55	13.12	12.88
LL CI (95%)	11.85	12.47	11.87

Prior to conducting the ANOVA, the assumption of normality was evaluated and determined to be satisfied as the three groups' distribution were associated with skew and kurtosis less then I2.0I and I9I , respectively.[14] (see Table 1.)

ANOVA test result pointed that the types of interview pattern preferred by the students for assessment of anatomy examination had no significant impact on obtained score **[F (2,240) = 2.80, p = .062].**

Furthermore to evaluate the nature of the differences between the three means further, the statistically non significant ANOVA was followed by three Pairwise independent sample t-test as a post-hoc verification with Bonferroni correction for significance level. Again there was no significant difference between groups. GROUP T vs. GROUP W (P=0.02), GROUP W vs. GROUP Z (P=0.14) and GROUP T vs. GROUP Z (P=0.58).

We are thus able to accept both hypothesis (H1 and H2)

Graph 1. Box and Whisker Plot
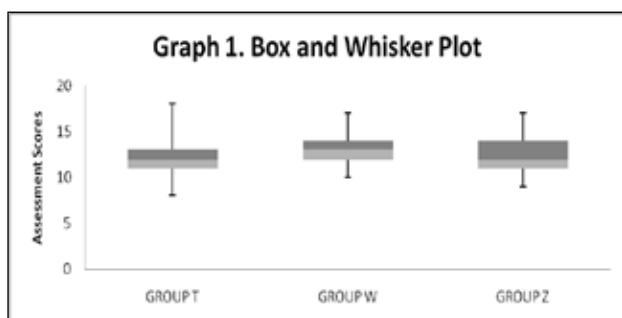
The median score with quartiles are reflected for each group in Box and Whisker Plot. Graph 1.

Statistically for all practical considerations overall score in each group were very close in nearly all three groups.

Discussion:
Based upon uncertainty in current scenario regarding regularization of academic session, we believed advantages of carrying interview rather than dropping assessment at present for future dates or extrapolate the marks obtained in previous assessment as generalization.[15,16] Inherently face to face interview has several strengths but is unfortunately impossible in current scenarion.[2,3,8,] e also wanted to survey the accomplishment and consistency in learning through ongoing audio-visuals classes in current disruption.[17] We, henceforth produce our observations based upon students viewpoints collected before and after the test through feedbacks.

- Preparation of suitable questionnaire: Preparing a pool of rational and communicative questionnaires takes consideration of variable responsiveness of students and obliteration of difficult explanatory questions. Questions tried to focus mainly on recollection of facts and its applicability. Framed questions were clearly structured and adaptable between selecting know/ don't know responses. Objective nature of questions with short answer was preferred to curtail the unfair control of interaction context by either of investigator or respondents.[18] A common pool of questionnaire was necessary to rule concerns of validity, consistency and variations at the end of interviewer. Assessor so far had been didactic in approach in framing questions on the spot while subtly measuring the caliber of candidates. It is entirely new experience for assessor to confine strictly to pre-decided pool of questions. Gradual acquisitions of experience in creating and working on such collections of questionnaire are highly desired. In feedback our students expressed pleasant satisfaction over the fact that interrogation would be from common questionnaire only.
- Preference for telephonic interview: surprisingly this was most dominant mode (n= 119/250: 4 were excluded in the final study design as discussed) chosen by students. Our observations obtained from feedbacks were:-
  - ➤ Familiarity of internet usage is distorted by age, gender and computer awareness.
  - ➤ We anticipated it to be most criticized method, but were taken by surprise by looking overwhelming response.
  - ➤ Easiest modality with random digital dialing suitable enough over all geographical areas was a reasonable determinant. In terms of accessibility,

nothing can beat the lowly but omnipresent telephone.

➢ Issue of reluctance of students for coming face to face to avoid interviewer bias in assessment of confidence.

➢ Interestingly, the numbers of missed questions per students per session was highest in this group.

➢ Few students in far flung geographical areas wished they could have fluent internet access and would have preferred other modalities. They admitted their inability to use catches in fine visual help provided by assessor was totally missing in telephonic interview. Many systematically ignore taking parts in telephone assessment for unknown reasons.

➢ The only input examiner receives is voice. Expressions and body language, produces a prejudice. The candidate is entitled to project confidence, knowledge and overall fit using voice only, thus creating limitations.

• Preference for Whatsapp video call interview: this was the second most preferred mode (n= 79/250: 2 were excluded in the final study design as discussed) after telephonic calls. The feedback for choosing this mode reflected:-

➢ Students didn't want to be left out as less smart when it comes face to face virtual interactive session.

➢ Students are convinced of displaying better responses, when eye contact is established.

➢ Whatsapp video call is an easy familiar technological platform and is ubiquitous for the group of students.

➢ A lot of students were not versed in exploring technicalities of cumbersome laptop based zoom meeting session and hence, preferred video call through whatsapp. These students purposefully denied Zoom platform meetings for problem pertaining to voice quality, video and basic connectivity which could pop up without warning, forcing to troubleshoot on technical problems rather than focus on assessment interview.

➢ Students didn't want to miss the opportunity of being able to be observed visually and create a short of interviewer bias.

➢ An internet access limitation in this group was not a major issue owing to urban background.

• Preference for Zoom platform meetings: relatively few students preferred this method (n= 45 /250: 1 was excluded in the final study design as discussed). Our observations as per feedback

were:

➢ They were heavy internet users registered on web platforms and have access to high speed internet and laptops.

➢ Typically these respondents expressed more discomfort about discussing topics over telephone then face to face.

➢ Also, this group had least numbers of missed questions per students per session.

➢ Interestingly, majority of students were in formal laboratory dresses; maybe they wanted to project voice and ideas.

➢ Non verbal cues in interaction were bright and better among three groups.

• Issues in assessment: Assessment works upon integration of dialogue and opinion exchanges.[17]Classically viva voce had been modality for evaluation in educational objectives acquired during self learning and training according to recommended curriculum.[19]Selection of assessment instrument must be congruent with students liking in changing scenario. No single assessment instrument is better overall as golden format, but we need to respect the core idea of near congruence with erstwhile time honored methodologies.[17] We are yet to discover blueprints in changing scenario through a long term cohort.[2] Since, the assessment was too objective, we missed out inclusion of demonstrative capabilities of all students. Assessment motivates students by group dynamics, enthusiasm of teachers. A positive outlook for assessment is expected of medical students and young doctors.[21,22] The additional noteworthy aspects of nontraditional discipline-independent skills (NTDIS) through their interaction are.[2,23]

➢ Presentation skills

➢ Oral communication skills

➢ Professional development skills

➢ IT competency skills

➢ Motivation to accept changes in pattern in line with the future.[24-26]

Conclusion:

We should not only investigate context based vs. content based knowledge of students, but also need to analyze the effectiveness of these modes. This study outline cross sectional study and compare the results of the analysis, highlighting the difference between selected modes. On a long term basis, we also need to test the three modes of assessment in terms of sensitive questions collected as feedback. However

advance technology might get, a complete simulation of conventional in-person interview in its full blown multisensory splendor is not achievable. The students and assessor are no doubt denied the interaction dynamics in seriousness of learning process. We need to see and maintain all assessment must be handled in the same way, have similar difficulties without distinction of any particular merit of examiner or students.[20] We need to further quantify reasons, explore evidences and convictions for approval of alternative promising futuristic assessment methods. In the background of inconclusive findings, it is fairly difficult for instructor to predict conclusively the benefit of either of three methods. Finally, in the conclusion, we give an overview of most important findings and discuss implications of further research. Further research can help instructor learn potential ways to tailor new variants of assessment methods.

Limitation of study

There were blind spots in interaction in all modes. Owing to variable factors, participants was not arbitrarily assigned to particular modality of assessment conditions, we were left with non-random generated sample data thus missing out the opportunity of randomization. Students were left freely to choose between three modalities of interview depending upon individual choices. The breadth of study is fairly limited being Cross-sectional and single- centric. Variable confounding entities ignored in performance during assessment might influence the outcomes. The involvement of nefarious cheating methods in particularly in telephonic interview, although minimized by time limitation of interviews, cannot be altogether denied.

Declaration for Ethical consideration: Above study had been designed after obtaining suitable clear consent from students for participation and feedbacks. Data were collected anonymously and confidentiality of the students' information was secured. The use of data for the purpose of study was communicated with ethical committee of the institute and the present work started only after clearance. The confidentiality of academic record of internal assessment marks was maintained and blinded to the statistician as per the direction of institutional ethical after obtaining approval.

Scope of further research and disclaimer

- The trends in changes of assessment method need to be evaluated in multicentre cohort
- Another limitation of the present study is that the selected types of validity, consistency and reliability of the questionnaire prepared for new methods of assessment should be analyzed. The authors are working on preparing the analysis following which instructors might be able to begin to address these related issues to help students in academic settings.
- We also need descriptive analysis of structured feedback collected through student's responses for anonymous questionnaires regarding new format of assessment. Authors are involved in additional explorations into the facts and data.
- *The constructed manuscript bears no conflicting interest amongst authors. Authorship priorities had been duly acknowledged. No financial implications were involved.*

References:

1. Ganguly PK, Chakravarty M, Latif NA, Osman M, Abu-Hijleh M. Teaching of anatomy in a problem-based curriculum at the Arabian Gulf University: the new face of the museum. *Clin Anat*. 2003;16(3):256-261. doi:10.1002/ca.10144.

2. Franchi T. The Impact of the Covid-19 Pandemic on Current Anatomy Education and Future Careers: A Student's Perspective. *Anat Sci Educ*. 2020;13(3):312-315. doi:10.1002/ase.1966.

3. Warner JH, Rizzolo LJ. 2006. Anatomical instruction and training for professionalism from the 19th to the 21st centuries. Clin Anat 19:403–414.

4. Smith CF, McManus B. 2015. The integrated anatomy practical paper: A robust assessment method for anatomy education today. AnatSciEduc 8:63–73.

5. Kumar N, Kanchan T, Unnikrishnan B, Thapar R, Mithra P, Kulkarni V, Holla R, Bhagwan D, Radhakrishnan Y. Incorporating problem based learning into medical curriculum: An experience from a medical college in Mangalore. Indian J Pharmacol. 2017 Sep-Oct;49(5):344-347. doi: 10.4103/ijp.IJP_492_16.

6. Langfield T, Colthorpe K, Ainscough L. 2018. Online instructional anatomy videos: Student usage, self-efficacy, and performance in upper limb regional anatomy assessment. AnatSciEduc 11:461–470.

7. O'Byrne PJ, Patry A, Carnegie JA. 2008. The development of interactive online learning tools for the study of anatomy. Med Teach 30:e260–e271.

8. Stone DM, Barry DS. 2019. Improving virtual learning interactions: Reducing the transactional distance of online anatomy modules. AnatSciEduc 12:686–687.

9. Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev. 2015 Sep 7;2015(9):CD002098.

10. Coleman E, O'Connor E. The role of WhatsApp® in medical education; a scoping review and instructional design model. BMC Med Educ. 2019 Jul 25;19(1):279. doi: 10.1186/s12909-019-1706-8.

11. Pather N, Blyth P, Chapman JA, Dayal MR, Flack NAMS, Fogg QA et al. Forced Disruption of Anatomy Education in Australia and New Zealand: An Acute Response to the Covid-19 Pandemic. Anat Sci Educ. 2020 May;13(3):284-300. doi: 10.1002/ase.1968.

12. Mishra, P., Singh, U., Pandey, C. M., Mishra, P., & Pandey, G. (2019). Application of student's *t*-test, analysis of variance, and covariance. *Annals of cardiac anaesthesia*, *22*(4), 407–411.

13. FigueiredoFilho, DalsonBritto, Paranhos, Ranulfo, Rocha, Enivaldo C. da, Batista, Mariana, Silva Jr., José Alexandre da, Santos, Manoel L. Wanderley D., & Marino, Jacira Guiro. (2013). When is statistical significance not significant?. *Brazilian Political Science Review*, *7*(1), 31-55.

14. Ho AD, Yu CC. Descriptive Statistics for Modern Test Score Distributions: Skewness, Kurtosis, Discreteness, and Ceiling Effects. EducPsychol Meas. 2015 Jun;75(3):365-388. doi: 10.1177/0013164414548576.

15. Collins JP. Modern approaches to teaching and learning anatomy. *BMJ*. 2008;337:a1310. Published 2008 Sep 9. doi:10.1136/bmj.a1310

16. Attardi SM, Choi S, Barnett J, Rogers KA. 2016. Mixed methods student evaluation of an online systemic human anatomy course with laboratory. AnatSciEduc 9:272–285.

17. Inuwa IM, Taranikanti V, Al-Rawahy M, Habbal O. 2012. Anatomy practical examinations: How does student performance on computerized evaluation compare with the traditional format? AnatSciEduc 5:27–32.

18. Turney BW. 2007. Anatomy in a modern medical curriculum. Ann R CollSurgEngl 89:104–107.

19. Erich Brenner, Andy R.M. Chirculescu, Concepción Reblet, and Claire Smith. Assessment in anatomy. Eur. J. Anat. 2015;19(1):105-124.

20. Choudhury B, Gouldsborough I, Shaw FL. 2016. The intelligent anatomy spotter: A new approach to incorporate higher levels of Bloom's taxonomy. AnatSciEduc 9:440–445.

21. Smith CF, Martinez-Álvarez C, McHanwell S. 2014. The context of learning anatomy: Does it make a difference? J Anat 224:270–278.

22. Louw G, Eizenberg N, Carmichael SW. The place of anatomy in medical education: AMEE Guide no 41. *Med Teach*. 2009;31(5):373-386.

23. Spampinato CM, Wittich CM, Beckman TJ, Cha SS, Pawlina W. "Safe Harbor": evaluation of a professionalism case discussion intervention for the gross anatomy course. *Anat Sci Educ*. 2014;7(3):191-198.

24. Estai M, Bunt S. 2016. Best teaching practices in anatomy education: A critical review. Ann Anat 208:151–157.

25. Erolin C, Reid L, McDougall S. 2019. Using virtual reality to complement and enhance anatomy education. J Vis Commun Med 42:93–101.

26. Evans DJ, Pawlina W, Lachman N. 2018. Human skills for human[istic] anatomy: An emphasis on nontraditional discipline-independent skills. AnatSciEduc 11:221–224.