

Redesigning an International Orthopaedic CME Course: The Effects on Participant Engagement over 5 Years

Abhiram R. Bhashyam^{a,b}, Quirine M.J. van der Vliet^b, R. Marijn Houwert^b, Rogier K. J. Simmermacher^b, Peter Brink^c, Piet de Boer^d and Luke P. H. Leenen^b

^aDepartment of Orthopaedic Surgery, Massachusetts General Hospital, Harvard Combined Orthopaedic Residency Program, Boston, MA, USA; ^bDepartment of Surgery, University Medical Centre Utrecht, Utrecht, The Netherlands; ^cDepartment of Traumatology, Maastricht University Medical Center, Maastricht, The Netherlands; ^dAO Foundation, Dubendorf, Switzerland

ABSTRACT

The time required to observe changes in participant evaluation of continuing medical education (CME) courses in surgical fields is unclear. We investigated the time required to observe changes in participant evaluation of an orthopaedic course after educational redesign using aggregate course-level data obtained from 1359 participants who attended one of 23 AO Davos Courses over a 5-year period between 2007 and 2011. Participants evaluated courses using two previously validated, 5-point Likert scales based on content and faculty performance, and we compared results between groups that underwent educational redesign incorporating serial needs assessment, problem-based learning, and faculty training initiatives (Masters Course), and those that did not (Non-Masters Course). Average scores for the usefulness and relevancy of a course and faculty performance were significantly higher for redesigned courses ($p < 0.0001$) and evaluations were significantly improved for both groups after faculty training was formalised in 2009 ($p < 0.001$). In summary, educational redesign incorporating serial needs assessment, problem-based learning, and faculty training initiatives were associated with improvement in participant evaluation, but these changes required 4–5 years to become evident.

ARTICLE HISTORY

Received 8 April 2019
Revised 3 June 2019
Accepted 12 June 2019

KEYWORDS

Continuing medical education; continuing professional development; needs assessment; orthopaedic surgery; training

Introduction

Continuing medical education (CME) is typically defined as an educational activity that maintains, develops, or increases the knowledge and skill a physician uses to provide services for patients, the public, or the profession once they have completed training [1,2]. Participation in CME activities is one way that physicians commit to continuous learning to provide the best evidence-based care, and CME is known to improve physician knowledge base, competence, and performance in practice [1,3–5]. CME is especially important for surgical education as new technology, techniques, and practice patterns can gradually lead to surgeon knowledge and performance gaps [6].

The AO Foundation is a medically guided non-profit organisation focused on the treatment of trauma and disorders of the musculoskeletal system, and it is one of the leading surgical CME providers for practising health-care professionals in orthopaedic trauma, veterinary, craniomaxillofacial, and spine surgery [6]. To deliver on its educational mission, the AO Foundation annually hosts a CME conference in Davos, Switzerland (“AO Courses Davos”). Within

this education programme, the AO Trauma Masters Courses in Davos are recognised as the flagship educational programme in orthopaedic trauma for the AO Foundation [7]. However, in 2005, AO Trauma Masters Courses were noted to have decreasing enrollment, poor participation and flagging enthusiasm. As a result, beginning in 2005, the AO Trauma Masters Courses underwent a process of structural redesign using the conceptual framework for CME later described by Moore, Green et al. [5,8] First, serial participant needs assessments were introduced [9], alongside a transition to case-based, problem-based learning (PBL) [10–12]. Subsequently, efforts to train the course faculty (“train the trainers” concept developed by Lisa Hadfield-Law) in gap analysis and facilitation of PBL for lectures and discussion groups were formalised (Table 1) [7,8,13,14]. Although these concepts are now used as a template by the AO, the outcomes of changes to the courses have only been studied in limited fashion [15,16].

In addition, changes to the AO Trauma Masters Courses took several years, but there are few long-term studies of educational redesign in surgical CME [17–22].

Table 1. AO Trauma Masters Course Changes.

Year	Change(s)
2005–2006	1. AO Learning Assessment Toolkit designed by Joseph Green piloted. (Standardised pre-and post-course needs assessment. Evaluation during the course of presentation relevancy and faculty performance) 2. Transition to problem-based learning
2007	Pilot programmes from 2006 are formalised.
2008	“Train the Trainers” approach initiated for AO Masters Courses. Expansion of the AO Learning Assessment Toolkit to other courses begins.
2009–2010	Formalised teacher training and blended online modules introduced for all orthopaedic trauma courses.

As a result, we have limited information about how long it takes for course changes to have an effect on participation and student satisfaction (the foundation required for successful CME programmes to enable learning, competence and performance) [5,9]. Thus, we explored the effect of the redesign of the AO Trauma Masters Courses (Davos) on participant evaluation of the usefulness and relevancy of course content and participant perception of faculty performance over a 5-year period (2007–2011). We hypothesised the course redesign would improve the evaluation of course relevancy and evaluation of faculty performance.

Methods

Ethical Consideration

Our study did not require IRB review as we used only anonymised aggregate course-level data without identifiable information about individuals.

Setting: The AO Davos Courses are an annual education event hosted in Davos, Switzerland in December. This CME conference has multiple courses in orthopaedic trauma, veterinary (VET), craniomaxillofacial (CMF), and spine surgery. The AO and the AO Davos Courses have been previously studied to understand orthopaedic education in other contexts [6,15,16,22].

Participants/Courses

Data were collected from surgeons attending the AO Davos Courses, and all responses were anonymised and reported in aggregate at the course level by the AO Foundation [15]. Participant enrollment by course and number of course offerings was recorded each year. In the primary analysis, the experimental group (Masters Course) consisted of AO Trauma Masters Courses which underwent an educational redesign. The control group (Non-Masters Course) consisted of other AO Trauma Non-Masters Courses targeted toward experienced orthopaedic providers, but which were not directly affected by educational redesign. We initially excluded courses directed towards trainees (e.g. AO Basic and

Advances courses), non-English language courses, and non-orthopaedic trauma courses (VET, CMF, and Spine). We included courses taught at the Davos conference between 2007 and 2011 and excluded results from pilot years (2006, 2008, and 2010). Each course had a lecture and discussion group component.

Outcome Measures

Electronic evaluation of course usefulness/relevancy and faculty performance by course participants was carried out in all courses using elements of the AO Learning Assessment Toolkit [16]. After each session during the course, participants were asked to evaluate each presentation (lecture or discussion group) using two previously validated questions based on content (“How useful and relevant was the content presented to your daily practice?”) and faculty performance (“How effective was the faculty in the role that he/she played?”) using a 5-point Likert scale (1 = not at all useful and relevant/effective, 5 = very useful and relevant/effective) [16]. Metrics on content and faculty performance were both used as a prior study demonstrated that accurate assessment of curriculum change requires simultaneous analysis of both factors [16].

Our study analysed course evaluation data that was aggregated at the course-level to protect data confidentiality concerns of participants and faculty by the AO Foundation. Each course had a lecture and discussion group component that was reported separately. We were provided with the course title, number of participants, minimum, mean, and maximum average Likert score for each course in a given year (for lecture and discussion groups).

Statistical Analysis

Cross-sectional descriptive statistics were calculated to quantify enrollment and participation by year. We performed a pooled analysis using weighted means by course participants to calculate aggregate mean Likert scores for each group (Masters and Non-Masters

Courses) [23–25]. Response range per group per year was calculated as the minimum average Likert score subtracted from the maximum average Likert score. We estimated the sample standard deviation based on number of participants using previously validated techniques [23]. We performed chi-square tests to assess for differences in response rate and compared differences in mean aggregate Likert score and response range using Welch's unequal variances *t*-test. Pearson correlation coefficient was used to assess the correlation between course relevancy and faculty performance evaluation and differences in response range for lectures versus discussion groups. We performed a sensitivity analysis using an alternative standard deviation estimation method and by including the AO Advanced Course within our Non-Masters Course control group (this course accepts advanced trainees/residents so was excluded in our main analysis) [25]. The significance criteria to assess for differences between groups were adjusted for multiple comparisons using the Bonferroni correction to $\alpha < 0.0031$. Stata software, version 14 (StataCorp), was used for all analyses.

Source of Funding

No funding source played a role in this investigation.

Results

Course Characteristics and Participation

We analysed 1359 participants who attended one of the 23 courses between 2007 and 2011 (Masters Courses 408 participants/6 courses; Non-Masters Courses 951

participants/17 courses). Detailed enrollment and the number of course offerings by year are presented in Figure 1.

Response Characteristics

Average response rate was 66% (range: 54%-74%). Response rate was similar within the Masters and Non-Masters Courses between 2007 and 2011, respectively. However, the average response rate was higher for Non-Masters Courses versus Masters Courses in 2007 (74% vs. 60%, $p = 0.004$) and 2011 (73% vs. 54%, $p = 0.001$). Average response range was 1.79 points on the Likert scale (range: 1.04–2.41). Response range for evaluation of course relevancy and faculty performance were correlated (correlation coefficient (r) = 0.88, $p = 0.021$) and similar between Masters and Non-Masters Courses ($p = 0.24$). The number of audience response questions (ARS) increased in both groups over the study period (Appendix Figure A1).

Lecture and Discussion Group Evaluation

For course lectures, average Likert score for lecture relevancy and faculty performance was significantly higher for Masters Courses compared to Non-Masters courses in 2007, 2009, and 2011 ($p < 0.0001$). Lecture evaluations for relevancy and performance for both groups were significantly higher in 2011 as compared to 2007 ($p < 0.0001$, Figure 2(a,b)). There was a very strong correlation between participants' perception of the usefulness and relevancy of the course to their practice and to their perception of faculty performance for

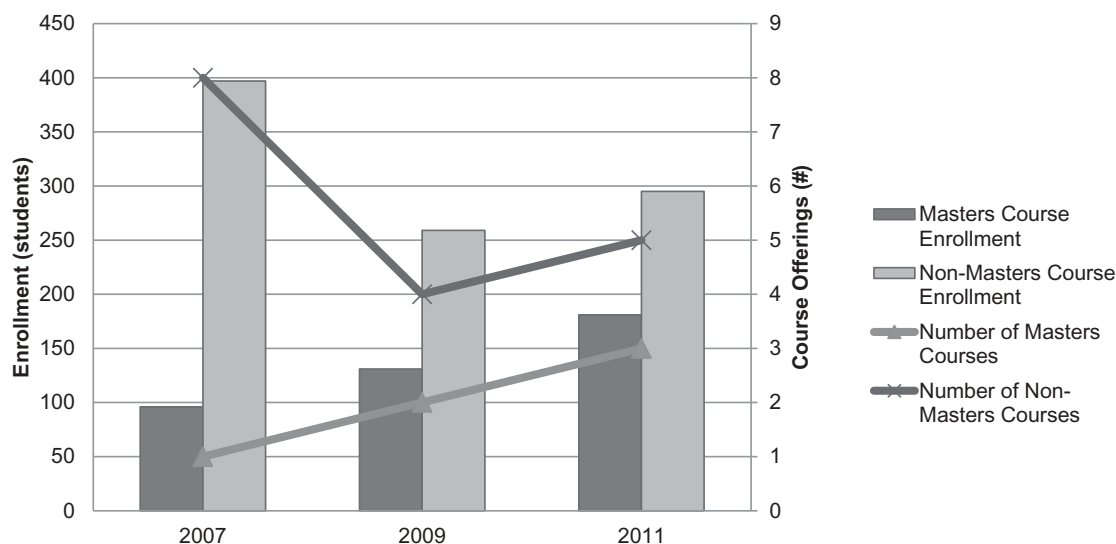


Figure 1. Course Participation.

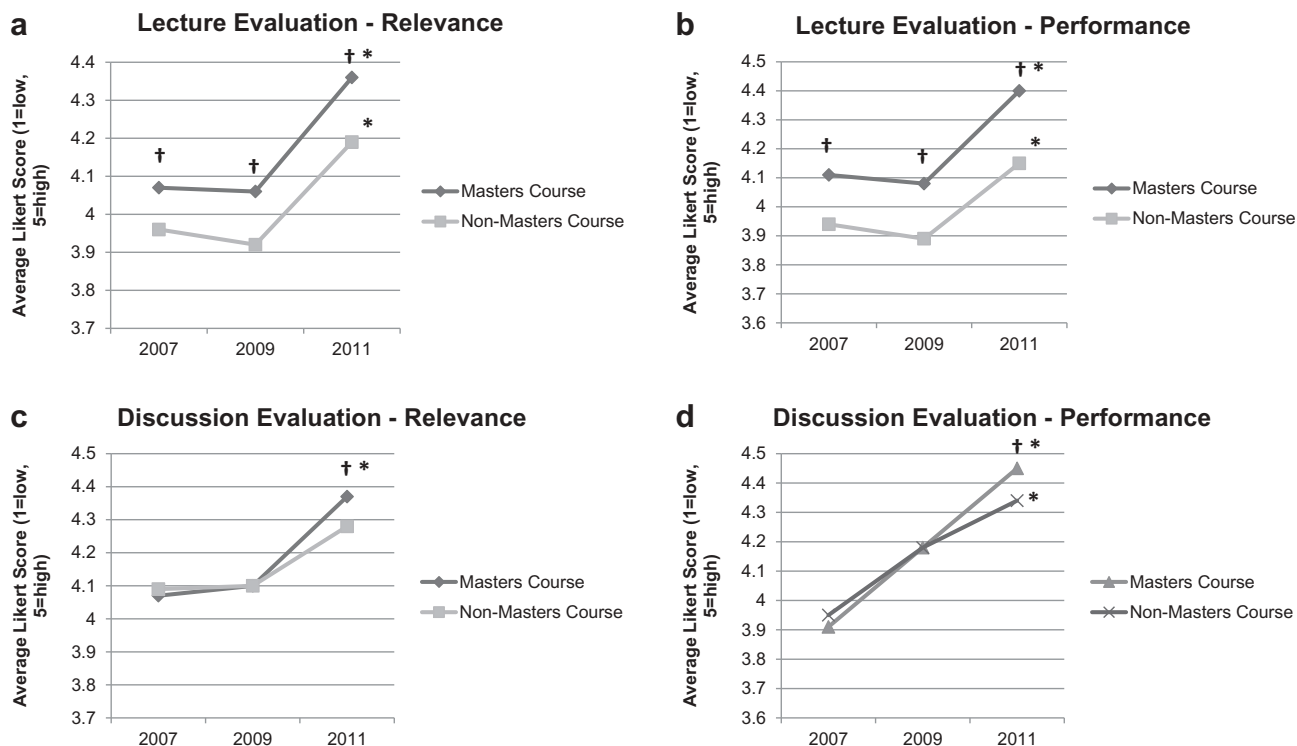


Figure 2. Course Evaluation. (a) Lecture Evaluation – Relevance; (b) Lecture Evaluation – Performance; (c) Discussion Evaluation – Relevance; (d) Discussion Evaluation – Performance; † = $p < 0.001$ (Masters Course vs. Non-Masters Course), * $p < 0.0001$ (2011 vs. 2007).

both the Masters Courses ($r = 0.998$, $p = 0.04$) and Non-Masters Courses ($r = 0.999$, $p = 0.028$).

For discussion group, average Likert score for discussion group relevancy and faculty performance was significantly higher for Masters Courses compared to Non-Masters Courses in 2011 ($p < 0.0001$). Evaluations for relevancy and faculty performance were similarly higher for both groups in 2011 compared to 2007 ($p < 0.0001$, Figure 2(c,d)). There was not a statistically significant correlation between topic relevancy and perception of faculty performance for the discussion groups (Masters Course, $r = 0.91$, $p = 0.275$; Non-Masters Course, $r = 0.84$, $p = 0.379$).

Sensitivity Analysis

Repeat analysis including Advanced courses with subspecialty courses led to identical results (Appendix Figure A2). Similar statistical results were obtained using a different estimation strategy for the standard deviation [25].

Discussion

Most physicians would agree that commitment to continuing learning is the best way to provide evidence-

based care, and CME is important in surgical education to remain abreast of new technology, techniques and practice patterns. In orthopaedic surgery, the preferred way to receive education in most countries is by non-industry-based courses [6,26]. Successful CME programmes require participation and satisfaction to enable learning, but there has been limited literature on how to increase participation and satisfaction within CME courses [5]. In addition, the time span required to observe changes after an intervention is unknown [17,19,21,22]. In this study of the redesign of the AO Trauma Masters Courses, we found that needs assessment, PBL, and faculty training initiatives led to significant improvement in participant evaluation of usefulness and relevancy of course content and participant perception of faculty performance. However, in some cases, these changes required 4–5 years to become evident.

Redesign of the AO Trauma Masters Courses used the conceptual framework, that was later elaborated upon by Moore et al., where CME development is viewed as a series of overlapping efforts that are part of a continuum linking lecture, discussion groups, and hands-on learning together [5]. Within this framework, three major changes were made to the Masters Courses to enable change (Figure 3). First, serial needs

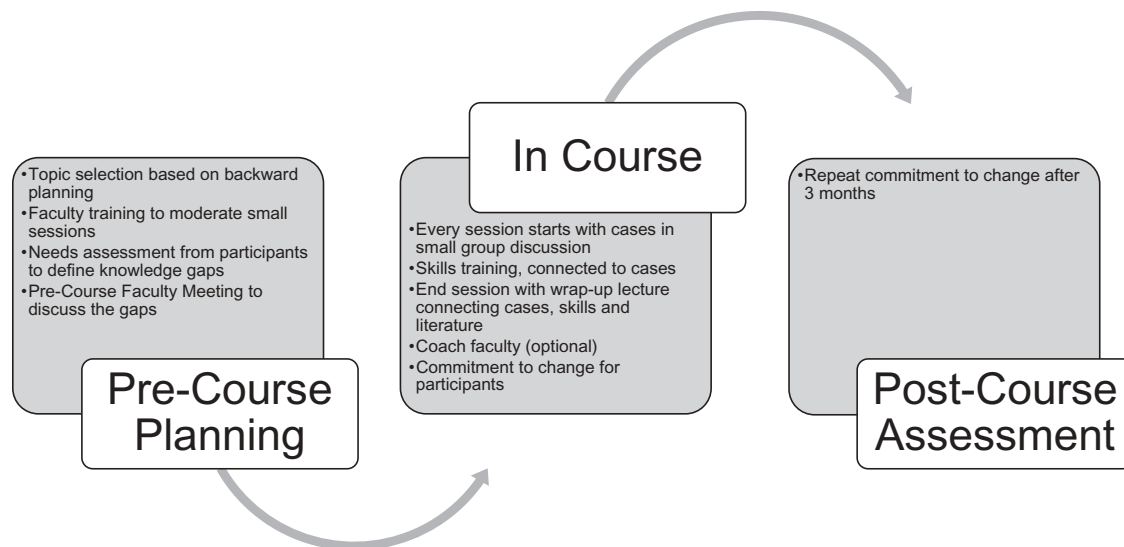


Figure 3. AO Masters Course Redesign Components and Execution.

assessments were implemented. A needs assessment is a powerful tool to discern an audience's desire for improvement and learning in CME activities which are more likely to lead to a change in practice when a needs assessment has been implemented [6,27,28]. Second, the course design was switched from lecture-based didactics to PBL in both lectures and discussion groups. Multiple prior studies have demonstrated that interactive techniques using case-based discussions allow for superior performance in improving the medical education environment [10–12,29]. Finally, faculty training initiatives (“train the trainers”) were formally implemented to help facilitate learning based on needs analysis and PBL concepts. The marked changes observed in participant perception of usefulness and relevancy of course content and their perception of faculty performance may be the result of synergy between these three elements of curricular redesign.

Our results also highlight the importance of faculty training in educational redesign efforts. While the AO Trauma Masters Course switched to PBL concepts in 2007, improvements in results were most significant in 2011 after faculty training efforts were fully rolled out in 2009. Faculty training was introduced in the UK by Lisa Hadfield-Law in 1990, started to become international in 2004, but was only formalised by 2009. Prior studies have demonstrated that the primary reason for participants to attend face-to-face courses was for interaction with expert faculty and discussion/feedback from experts, while the majority of course attendants changed their opinion to the correct, evidence-based option after case-based discussion driven by experts [6,17]. Similarly, we saw the largest improvements in participant evaluation after

faculty training efforts. In addition, evaluation of course relevancy was strongly correlated with faculty performance. The importance of faculty in enabling learning in CME courses cannot be overemphasised.

Finally, multiple prior studies have shown the short-term effects of the introduction of new CME conferences or courses, but there has been limited information on expected changes beyond a single course or year [17,18,21,22]. Our results highlight that three or more years are sometimes required to see the effects of the educational redesign at the course-level. This is especially important at a policy level where funding agencies often expect changes at a much more rapid pace [32]. While the AO Foundation is large and can afford to make changes over longer time spans, changes on this time scale many not be tenable for smaller educational efforts and their funders (e.g. low-middle income countries) [18]. As in other settings, our findings emphasise the need for patience to withstand the lag time between change initiation and result observation [32].

Limitations

Our study has several limitations. Our analysis was limited to English-language courses at a single site in Europe, although this is the “gold-standard” course for most orthopaedic surgeons. Our response rate to evaluation questions varied between 54% and 74% allowing for possible response bias. However, this response rate is common for this type of research and is much improved compared to other studies with similar context [15,18,33]. In addition, there may have been some element of survey fatigue as

average response rate was higher for Non-Masters Courses versus Masters Courses in 2007–2011, and, in those years, the number of survey questions was significantly higher for the Masters Courses (Appendix Figure A1). One major limitation of this analysis is that we only assessed participant satisfaction/engagement. However, student engagement is important to facilitate learning and future studies may focus on the longitudinal effects of curricular change on learning outcomes [5,9]. Future work should also focus on longitudinal assessments of learning using techniques like gap analysis. The minimum significant difference for changes in average Likert scale scores for this context has not previously been defined. However, we considered it significant that average scores increased by 0.2–0.3 points – corresponding to an increase of 5–10% compared to 2007 values or 15–20% of the average response range. Our investigation employed pooled analysis of course-level data because of data confidentiality concerns associated with using individual data; however, this approach is well described in the current literature. Care should be exercised when employing the results of this study to individuals given the risk of committing the ecological fallacy [34,35]. Finally, in using pooled aggregate data analysis, we estimated the standard deviation for statistical analysis using validated techniques, but it is possible that different results could be obtained when using individual data. However, other studies have used a similar setting and approach [15,16]. Our results were unchanged when using another standard deviation estimation strategy [25]. In addition, others have shown that analysis of summary information yields similar results as when using individual data [24].

Conclusions

Our results are of relevance to physicians and educators interested in surgical CME. We demonstrate how the introduction of serial needs assessment, problem-based learning, and faculty training initiatives to the flagship AO Trauma Masters Course led to significant improvement in participant evaluation of course content and participant perception of faculty performance. However, in some cases, these changes required 4–5 years to become evident, and the largest changes were only seen after faculty training initiatives. Systematic change to educational programmes can improve participant engagement allowing for subsequent development of learning, competence and performance [5]. This philosophy may be able to improve the alignment of surgeon practice with evidence-based guidelines that lead to better patient care, but this requires further study [15,36].

Significance

Introduction of serial needs assessment, problem-based learning, and faculty training initiatives improved participant evaluation of course content and faculty performance. In some cases, these changes required 4–5 years to become evident, and the greatest improvement was observed after the introduction of faculty training initiatives.

Disclosure statement

No declared conflicts of interest. No external funding source.

References

- [1] Davis D. Continuing medical education. Global health, global learning. *BMJ*. 1998;316(7128):385–389.
- [2] AHRQ. *Effectiveness of Continuing Medical Education*.
- [3] Haynes RB, Davis DA, McKibbon A, et al. A critical appraisal of the efficacy of continuing medical education. *JAMA*. 1984;251:61–64.
- [4] Davis NL, Willis CE. A new metric for continuing medical education credit. *J Contin Educ Health Prof*. 2004;24:139–144.
- [5] Moore DE, Green JS, Gallis HA. Achieving desired results and improved outcomes: integrating planning and assessment throughout learning activities. *J Contin Educ Health Prof*. 2009;29:1–15.
- [6] Buckley R, Brink P, Kojima K, et al. International needs analysis in orthopaedic trauma for practising surgeons with a 3-year review of resulting actions. *J Eur CME*. 2017;6:1398555.
- [7] AO Foundation. *Annual Report 2006*. 2006.
- [8] AO Foundation. *Annual Report 2007*. 2007.
- [9] Hak DJ. Changing nature of continuing medical education. *Orthopedics*. 2013;36:748.
- [10] Bower EA, Girard DE, Wessel K, et al. Barriers to innovation in continuing medical education. *J Contin Educ Health Prof*. 2008;28:148–156.
- [11] Salti IS. Continuing medical education. *Med Educ*. 1995;29(Suppl 1):97–99.
- [12] Stancic N, Mullen PD, Prokhorov AV, et al. Continuing medical education: what delivery format do physicians prefer? *J Contin Educ Health Prof*. 2003;23:162–167.
- [13] AO Foundation. *Annual Report 2008*. 2008.
- [14] AO Foundation. *Annual Report 2010*.
- [15] de Boer PG, Buckley R, Schmidt P, et al. Barriers to orthopaedic practice—why surgeons do not put into practice what they have learned in educational events. *Injury*. 2011;43:290–294.
- [16] de Boer PG, Buckley R, Schmidt P, et al. Learning assessment toolkit. *J Bone Joint Surg Am*. 2010;92:1325–1329.
- [17] Berjano P, Villafaña JH, Vanacker G, et al. The effect of case-based discussion of topics with experts on learners' opinions: implications for spinal education and training. *Eur Spine J*. 2017;27(Suppl 1):2–7.

- [18] Fils J, Bhashyam AR, Pierre JBP, et al. Short-term performance improvement of a continuing medical education program in a low-income country. *World J Surg.* 2015;39(10):2407–2412.
- [19] Peter NA, Pandit H, Le G, et al. Delivering a sustainable trauma management training programme tailored for low-resource settings in East, Central and Southern African countries using a cascading course model. *Injury.* 2015;47:1128–1134.
- [20] Ali J, Adam R, Stedman M, et al. Cognitive and attitudinal impact of the advanced trauma life support program in a developing country. *J Trauma.* 1994;36(5):695–702.
- [21] Kelly M, Bennett D, Bruce-Brand R, et al. One week with the experts: a short course improves musculoskeletal undergraduate medical education. *J Bone Joint Surg Am.* 2014;96:e39.
- [22] Egol KA, Phillips D, Vongbandith T, et al. Do orthopaedic fracture skills courses improve resident performance? *Injury.* 2014;46:547–551.
- [23] Wan X, Wang W, Liu J, et al. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol.* 2014;14:135.
- [24] Moineddin R, Urquia ML. Regression analysis of aggregate continuous data. *Epidemiology.* 2014;25:929–930.
- [25] Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol.* 2005;5:13.
- [26] Schulz AP, Jönsson A, Kasch R, et al. Sources of information influencing decision-making in orthopaedic surgery - an international online survey of 1147 orthopaedic surgeons. *BMC Musculoskelet Disord.* 2013;14:96.
- [27] Hoyal FM. Skills and topics in continuing medical education for rural doctors. *J Contin Educ Health Prof.* 2001;20:13–19.
- [28] Qin Y, Wang Y, Floden RE. The effect of problem-based learning on improvement of the medical educational environment: a systematic review and meta-analysis. *Med Princ Pract.* 2016;25(6):525–532.
- [29] Sun M, Loeb S, Grissom JA. Building teacher teams: evidence of positive spillovers from more effective colleagues. *Educ Eval Policy Anal.* 2017;39:104.
- [30] Dodgson M, Gann D The missing ingredient in innovation: patience. <https://wwweforumorg/agenda/2018/04/patient-capital>. 2018.
- [31] Bhashyam AR, Fils J, Lowell J, et al. A novel approach for needs assessment to build global orthopedic surgical capacity in a low-income country. *J Surg Educ.* 2015;72. DOI:10.1016/j.jsurg.2014.10.008
- [32] Pearce N. The ecological fallacy strikes back. *J Epidemiol Community Health.* 2000;54(5):326–327.
- [33] Greenland S, Robins J. Invited commentary: ecologic studies—biases, misconceptions, and counterexamples. *Am J Epidemiol.* 1994;139(8):747–760.
- [34] Neily J, Mills PD, Young-Xu Y, et al. Association between implementation of a medical team training program and surgical mortality. *JAMA.* 2010;304:1693–1700.

Appendix

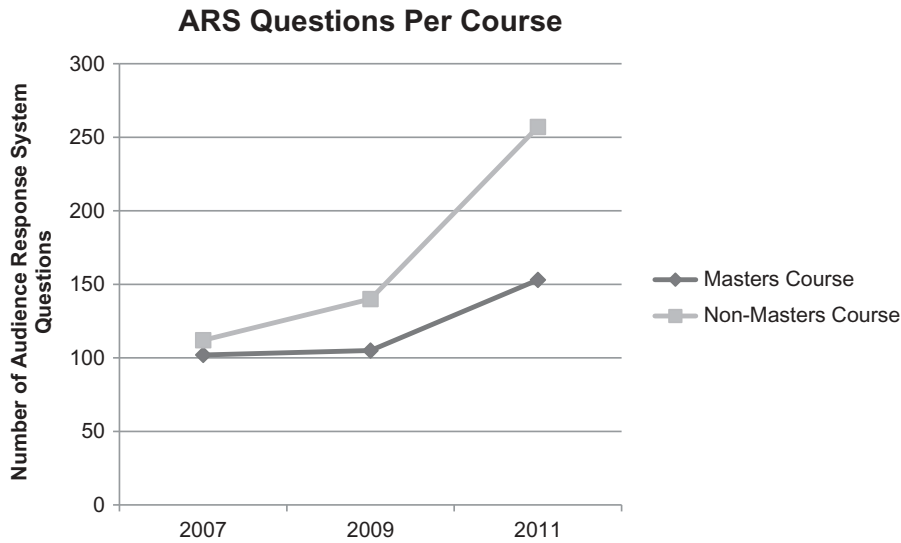


Figure A1. Audience Response System (ARS) Questions per Course.

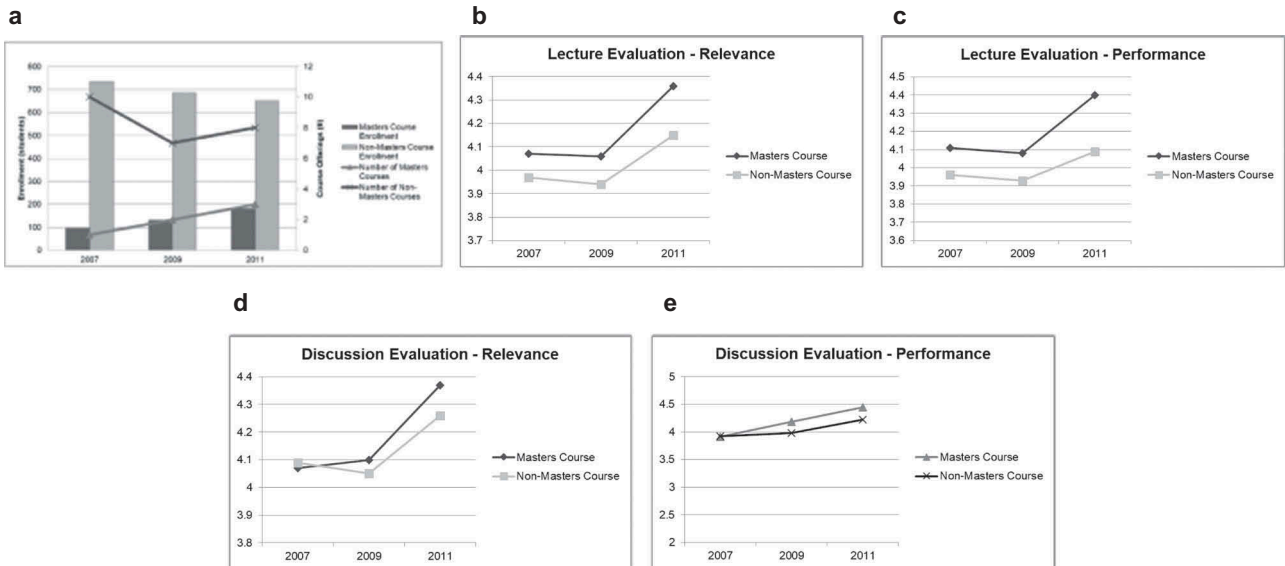


Figure A2. Sensitivity Analysis including AO Advanced Courses within Non-Masters Courses. (a) Course Participation; (b) Lecture Evaluation – Relevance; (c) Lecture Evaluation – Performance; (d) Discussion Evaluation – Relevance; (e) Discussion Evaluation – Performance; VET/CMF = Veterinary/Craniomaxillofacial Surgery Courses.