Feedback in formative OSCEs: comparison between direct observation and video-based formats

JUNOD PERRON, Noëlle Astrid, et al.

Abstract

Medical students at the Faculty of Medicine, University of Geneva, Switzerland, have the opportunity to practice clinical skills with simulated patients during formative sessions in preparation for clerkships. These sessions are given in two formats: 1) direct observation of an encounter followed by verbal feedback (direct feedback) and 2) subsequent review of the videotaped encounter by both student and supervisor (video-based feedback). The aim of the study was to evaluate whether content and process of feedback differed between both formats.

Reference


PMID : 27834170
DOI : 10.3402/meo.v21.32160

Available at:
http://archive-ouverte.unige.ch/unige:90303

Disclaimer: layout of this document may differ from the published version.
Feedback in formative OSCEs: comparison between direct observation and video-based formats

Noëlle Junod Perron1,2*, Martine Louis-Simonet3, Bernard Cerutti1, Eva Pfarrwaller4, Johanna Sommer4 and Mathieu Nendaz1,3

1Unit of Development and Research in Medical Education, Faculty of Medicine, University of Geneva, Geneva, Switzerland; 2Division of Primary Care Medicine, Department of Community Medicine, Primary Care and Emergencies, Geneva University Hospitals, Geneva, Switzerland; 3Service of General Internal Medicine, Department of General Internal Medicine, Rehabilitation and Geriatric Medicine, Geneva University Hospitals, Geneva, Switzerland; 4Unit of Primary Care Teaching and Research, Faculty of Medicine, University of Geneva, Geneva, Switzerland

Introduction: Medical students at the Faculty of Medicine, University of Geneva, Switzerland, have the opportunity to practice clinical skills with simulated patients during formative sessions in preparation for clerkships. These sessions are given in two formats: 1) direct observation of an encounter followed by verbal feedback (direct feedback) and 2) subsequent review of the videotaped encounter by both student and supervisor (video-based feedback). The aim of the study was to evaluate whether content and process of feedback differed between both formats.

Methods: In 2013, all second- and third-year medical students and clinical supervisors involved in formative sessions were asked to take part in the study. A sample of audiotaped feedback sessions involving supervisors who gave feedback in both formats were analyzed (content and process of the feedback) using a 21-item feedback scale.

Results: Forty-eight audiotaped feedback sessions involving 12 supervisors were analyzed (2 direct and 2 video-based sessions per supervisor). When adjusted for the length of feedback, there were significant differences in terms of content and process between both formats; the number of communication skills and clinical reasoning items addressed were higher in the video-based format (11.29 vs. 7.71, p = 0.002 and 3.71 vs. 2.04, p = 0.010, respectively). Supervisors engaged students more actively during the video-based sessions than during direct feedback sessions (self-assessment: 4.00 vs. 3.17, p = 0.007; active problem-solving: 3.92 vs. 3.42, p = 0.009). Students made similar observations and tended to consider that the video feedback was more useful for improving some clinical skills.

Conclusion: Video-based feedback facilitates discussion of clinical reasoning, communication, and professionalism issues while at the same time actively engaging students. Different time and conceptual frameworks may explain observed differences. The choice of feedback format should depend on the educational goal.

Keywords: formative; OSCEs; feedback; quality; video review; direct observation

Responsible Editor: Lynn Yeoman, Baylor College of Medicine, USA.

*Correspondence to: Noëlle Junod Perron, Division of Primary Care Medicine, Department of Community Medicine, Primary Care and Emergencies, Geneva University Hospitals, 4 rue Gabrielle Perret-Gentil, CH-1211, Geneva 14, Switzerland, Email: noelle.junod@hcuge.ch

Received: 6 May 2016; Revised: 19 September 2016; Accepted: 21 September 2016; Published: 8 November 2016
The aim of the study was to evaluate differences in the content and process of feedback in direct observation versus video-based OSCE formats.

Methods

Design, setting and participants
We conducted the study at the Faculty of Medicine, University of Geneva, Switzerland. The medical school offers a 6-year curriculum divided into three preclinical years (bachelor) and three clinical years (master) to 150 medical students per class. During the second and third preclinical years, medical students practice clinical skills during formative OSCEs in two different formats:

1. A 20-min interaction with a simulated patient followed by a 15-min feedback (direct observation) by a supervisor and 2–3 min of feedback by the simulated patient.
2. A 20-min videotaped interaction with a simulated patient, followed by immediate feedback by the simulated patient and a delayed 40-min feedback session by the supervisor 1 week later (including observation of the videotaped consultation), after the student has reviewed and analyzed his own performance (video-based feedback).

Data collection
During the 2012–2013 academic year, all feedback sessions given to second- and third-year medical students during formative OSCEs were audiotaped. For each supervisor who had given both types of feedback, two feedback sessions given in the ‘direct observation’ format and two feedback sessions given in the ‘video-based format’ were randomly selected.

Instrument
We adapted a previously developed instrument to assess the effectiveness of training on faculty feedback skills (11). The instrument, which assessed the way feedback is given, follows the structure of the MAAS-Global-Score, a well-known communication skills coding instrument, including both specific items, five global dimensions and a global rating item (12). Content validity was insured by an extensive review of the literature on effective feedback (5, 13–17). Construct validity was previously demonstrated in the ability of the instrument to detect improvements in faculty feedback skills (11). We added four items related to the content derived from the OSCE checklist used by supervisors and included three additional items specifically evaluating clinical reasoning and communication/professionalism.

Outcome measures

Content of the feedback
Content analysis consisted of identifying and quantifying all themes mentioned and/or discussed during the feedback session. Themes were divided into seven categories (Table 1): global, history taking, physical examination, content of the explanation-end of encounter, communication, elaboration on clinical reasoning, and elaboration on communication and professionalism. ‘Elaboration’ referred to the number of times the supervisor elaborated in a directive or facilitative way on the importance of collecting elements of the history taking or physical examination in relation to the clinical reasoning process or underlined the importance of communication skills or attitudes for future practice during the feedback session.

Process of the feedback
The process of feedback, focusing on the structure and the content of teaching skills used by the supervisors, was coded by N.J.P. and M.L.S. according to the 14 themes displayed in Table 1, using a 0–5 Likert scale (0 being absent and 5 optimal).

Interrater reliability, which was calculated on the basis of 10% of the audiotaped feedback sessions, was good (intraclass correlation coefficient = 0.89).

In addition, all students who received feedbacks from supervisors involved in both formats were also asked to rate the utility and quality of the feedback immediately after the feedback session. The self-administrated questionnaire consisted of 14 items used in a previous study (Table 2) (Junod Perron et al).

The overall project was approved by the research ethics committee of the University Hospitals of Geneva. A complete review was waived by this committee.

Analysis
Sociodemographic data were summarized by percentages, means, and standard deviation (SD). Feedback content data were expressed as the mean number of items
addressed per session of feedback. Feedback process data were summarized by mean scores of Likert scale and SD. Comparisons of feedback scores (for both feedback content and process) were made with multivariate analysis of variance including a supervisor effect, and a feedback duration effect categorized in four groups (the quartiles were chosen as cut-off values). Differences in students’ perceptions regarding the quality of feedback between both formats were analyzed using the ANOVA test. All analyses were run on R 2.15.3 (the R Foundation for Statistical Computing), and TIBCO Spotfire S+® 8.1 for Windows (TIBCO Software Inc).

Results
Twelve supervisors were involved in both direct observation and video-based feedback sessions. For each supervisor, two feedback sessions per format were randomly selected and analyzed (n = 48).

Table 1. Content of the feedback given by the supervisors according to two types of formative OSCE (12 supervisors, 48 audiotaped feedback)

<table>
<thead>
<tr>
<th>Number of items addressed/discussed</th>
<th>Video-based, mean (SD)</th>
<th>Direct observation, mean (SD)</th>
<th>Delta (SE)</th>
<th>p</th>
<th>Adjusted to the length of the feedback, Delta (SE)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global comments</td>
<td>0.50 (0.66)</td>
<td>0.63 (0.58)</td>
<td>−0.13 (0.19)</td>
<td>0.506</td>
<td>0.35 (0.35)</td>
<td>0.506</td>
</tr>
<tr>
<td>History taking</td>
<td>4.29 (3.21)</td>
<td>5.63 (3.21)</td>
<td>−1.33 (0.78)</td>
<td>0.097</td>
<td>−3.89 (1.38)</td>
<td>0.076</td>
</tr>
<tr>
<td>Physical exam</td>
<td>6.38 (2.78)</td>
<td>5.88 (2.51)</td>
<td>0.50 (0.72)</td>
<td>0.490</td>
<td>−1.12 (1.31)</td>
<td>0.476</td>
</tr>
<tr>
<td>Explanation – end of encounter</td>
<td>1.25 (1.19)</td>
<td>1.00 (1.06)</td>
<td>0.25 (0.32)</td>
<td>0.441</td>
<td>−0.66 (0.58)</td>
<td>0.416</td>
</tr>
<tr>
<td>Communication skills</td>
<td>11.29 (4.58)</td>
<td>7.71 (4.85)</td>
<td>3.58 (1.07)</td>
<td>0.002</td>
<td>1.05 (2.01)</td>
<td>0.002</td>
</tr>
<tr>
<td>Elaboration on clinical reasoning</td>
<td>3.71 (2.97)</td>
<td>2.04 (1.49)</td>
<td>1.67 (0.61)</td>
<td>0.010</td>
<td>−0.53 (1.10)</td>
<td>0.007</td>
</tr>
<tr>
<td>Elaboration on communication – professionalism</td>
<td>4.29 (1.68)</td>
<td>2.21 (1.38)</td>
<td>2.08 (0.40)</td>
<td>&lt;0.001</td>
<td>1.12 (0.76)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 2. Process of the feedback given by the supervisors according to two types of formative OSCE (12 supervisors, 48 audiotaped feedback)

<table>
<thead>
<tr>
<th>Feedback processa</th>
<th>Video-based, mean (SD)</th>
<th>Direct observation, mean (SD)</th>
<th>Delta (SE)</th>
<th>p</th>
<th>Adjusted to the length of the feedback, delta (SE)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>The supervisor explored students’ learning needs</td>
<td>3.75 (1.07)</td>
<td>3.25 (1.19)</td>
<td>0.50 (0.29)</td>
<td>0.090</td>
<td>0.30 (0.57)</td>
<td>0.104</td>
</tr>
<tr>
<td>The supervisor stimulated students’ self-assessment</td>
<td>4.00 (1.32)</td>
<td>3.17 (1.17)</td>
<td>0.83 (0.29)</td>
<td>0.007</td>
<td>1.01 (0.57)</td>
<td>0.009</td>
</tr>
<tr>
<td>The feedback was descriptive</td>
<td>4.33 (0.96)</td>
<td>4.17 (0.82)</td>
<td>0.17 (0.21)</td>
<td>0.422</td>
<td>−0.08 (0.40)</td>
<td>0.433</td>
</tr>
<tr>
<td>The feedback was subjective</td>
<td>4.46 (1.18)</td>
<td>3.91 (1.35)</td>
<td>0.54 (0.28)</td>
<td>0.059</td>
<td>0.10 (0.54)</td>
<td>0.066</td>
</tr>
<tr>
<td>The feedback was balanced</td>
<td>3.92 (1.18)</td>
<td>4.21 (1.10)</td>
<td>−0.29 (0.28)</td>
<td>0.303</td>
<td>0.02 (0.51)</td>
<td>0.287</td>
</tr>
<tr>
<td>The supervisor took into account the student's self-assessment</td>
<td>3.50 (1.59)</td>
<td>3.00 (1.35)</td>
<td>0.50 (0.35)</td>
<td>0.158</td>
<td>0.45 (0.68)</td>
<td>0.170</td>
</tr>
<tr>
<td>The supervisor stimulated the student to participate in the problem-solving process</td>
<td>3.92 (0.78)</td>
<td>3.42 (0.83)</td>
<td>0.50 (0.19)</td>
<td>0.012</td>
<td>−0.22 (0.34)</td>
<td>0.009</td>
</tr>
<tr>
<td>The supervisor used role-playing or hands-on practice to give students the opportunity to practice</td>
<td>1.75 (1.36)</td>
<td>1.46 (1.28)</td>
<td>0.29 (0.29)</td>
<td>0.321</td>
<td>−0.70 (0.53)</td>
<td>0.305</td>
</tr>
<tr>
<td>The supervisor checked the student’s understanding at the end of the feedback</td>
<td>3.79 (1.10)</td>
<td>3.25 (1.48)</td>
<td>0.54 (0.30)</td>
<td>0.083</td>
<td>0.35 (0.55)</td>
<td>0.069</td>
</tr>
<tr>
<td>Empathy</td>
<td>4.13 (0.74)</td>
<td>4.17 (1.13)</td>
<td>−0.04 (0.19)</td>
<td>0.825</td>
<td>−0.23 (0.37)</td>
<td>0.832</td>
</tr>
<tr>
<td>Pedagogical effectiveness</td>
<td>3.92 (0.72)</td>
<td>3.71 (0.91)</td>
<td>0.21 (0.17)</td>
<td>0.221</td>
<td>−0.17 (0.31)</td>
<td>0.216</td>
</tr>
<tr>
<td>Structure of the feedback</td>
<td>4.42 (0.83)</td>
<td>3.88 (0.95)</td>
<td>0.54 (0.16)</td>
<td>0.001</td>
<td>0.45 (0.30)</td>
<td>0.002</td>
</tr>
<tr>
<td>Verbal interaction</td>
<td>4.08 (0.58)</td>
<td>3.54 (1.06)</td>
<td>0.54 (0.19)</td>
<td>0.007</td>
<td>−0.08 (0.35)</td>
<td>0.007</td>
</tr>
<tr>
<td>Global evaluation</td>
<td>4.19 (0.55)</td>
<td>3.88 (0.78)</td>
<td>0.31 (0.12)</td>
<td>0.012</td>
<td>−0.03 (0.21)</td>
<td>0.009</td>
</tr>
</tbody>
</table>

aLikert scale 0–5: 0 being absent and 5 optimal.
Supervisors included seven women (7/12, 58%) and supervisors’ mean age was 45 years old (SD 10). Three worked in hospital-based general internal medicine, four in ambulatory general internal medicine, three as educationalists, and one as a hospital specialist. They had on average 18.2 years (SD 9.9) of clinical experience, had been clinical supervisors for an average of 9.4 years (SD 8.7), and were involved as supervisors in the formative OSCE for 4.8 years (SD 5.7).

Mean length of feedback was 26.2 min (SD 7.9) for the video-based sessions and 13.5 min (SD 4.0) for the direct observation sessions ($p < 0.001$). Length of feedback in video-based sessions was measured as the actual time of verbal interaction, excluding video observation time.

During videotaped feedback sessions, supervisors addressed more often communication issues and elaborated more often on clinical reasoning and communication and professionalism issues than during direct observation sessions, independent of the length of feedback (Table 1). Supervisors were also more learner centered during video-based feedback sessions (Table 2) and scored higher for several elements such as students’ self-assessment, student’s participation in the problem-solving process, and checking of students’ understanding, and transversal dimensions such as feedback structure and the amount of verbal interaction.

The 140 students who received feedback from supervisors involved in both formats perceived that in the video-based feedback, the supervisor made them more active in the problem-solving. There was also some evidence that the video format was more effective than the direct observation format at improving their physical exam and communication skills (Table 3).

### Table 3. Students’ perceptions of the quality of feedback according to the type of feedback format (12 supervisors and 140 students)

<table>
<thead>
<tr>
<th>Quality of the feedback perceived by the students$^{a}$</th>
<th>Video-based $n = 109$, mean (SD)</th>
<th>Direct observation $n = 42$, mean (SD)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>The feedback session was useful</td>
<td>4.81 (0.46)</td>
<td>4.75 (0.57)</td>
<td>0.529</td>
</tr>
<tr>
<td>I improved my history taking skills</td>
<td>4.43 (0.74)</td>
<td>4.40 (0.75)</td>
<td>0.813</td>
</tr>
<tr>
<td>I improved my physical examination skills</td>
<td>4.51 (0.87)</td>
<td>4.20 (0.94)</td>
<td>0.063</td>
</tr>
<tr>
<td>I improved my communication skills</td>
<td>4.49 (0.71)</td>
<td>4.19 (0.90)</td>
<td>0.060</td>
</tr>
<tr>
<td>The supervisor was aware of what I needed to learn</td>
<td>4.90 (0.30)</td>
<td>4.82 (0.43)</td>
<td>0.269</td>
</tr>
<tr>
<td>The supervisor made me feel comfortable and confident</td>
<td>4.80 (0.46)</td>
<td>4.63 (0.67)</td>
<td>0.137</td>
</tr>
<tr>
<td>The supervisor asked me my learning needs</td>
<td>4.48 (0.96)</td>
<td>3.62 (1.55)</td>
<td>0.001</td>
</tr>
<tr>
<td>The supervisor asked me to evaluate what I did well</td>
<td>4.69 (0.52)</td>
<td>4.46 (0.91)</td>
<td>0.127</td>
</tr>
<tr>
<td>The supervisor asked me to evaluate what I could improve</td>
<td>4.93 (0.26)</td>
<td>4.77 (0.53)</td>
<td>0.067</td>
</tr>
<tr>
<td>The supervisor gave me balanced feedback (including both positive and less positive aspects)</td>
<td>4.93 (0.26)</td>
<td>4.69 (0.64)</td>
<td>0.022</td>
</tr>
<tr>
<td>The supervisor stimulated me to participate in the problem-solving process</td>
<td>4.56 (0.64)</td>
<td>4.00 (1.09)</td>
<td>0.003</td>
</tr>
<tr>
<td>The supervisor gave me precise and concrete suggestions for improvement</td>
<td>4.71 (0.64)</td>
<td>4.44 (0.90)</td>
<td>0.072</td>
</tr>
<tr>
<td>The supervisor provided me opportunities to practice parts of the history taking, physical exam, or the communication</td>
<td>3.18 (1.52)</td>
<td>2.77 (1.58)</td>
<td>0.165</td>
</tr>
<tr>
<td>The supervisor checked my understanding</td>
<td>3.58 (1.48)</td>
<td>3.37 (1.56)</td>
<td>0.473</td>
</tr>
</tbody>
</table>

$^{a}$Likert scale 1–5: 1 completely disagree and 5 completely agree.

Discussion

The results show that supervisors gave different feedback both in terms of content and process during video-based or direct observation formats, independent of the duration of feedback. In the video-based format, supervisors are more learner centered, address more communication and professionalism issues, and make more active links between history taking and physical examination and clinical reasoning. Students also perceived that the video format was more learner centered, and there was also some evidence that it was more effective at improving their communication and physical exam skills.

We found no literature about how the format in which formative OSCE takes place influences the way supervisors give feedback. In the field of simulation, the use of video-assisted debriefing is not superior to the use of non-video-assisted debriefing for the acquisition of technical and non-technical skills (18). However, regarding non-technical skills, as stated in two reviews, no study specifically analyzed the impact of the debriefing structure on learners’ perceptions and performance (19, 20).

Several factors may explain such differences of feedback formats. Direct observation and video-based formats use different frames of time. The first takes place immediately after the observed performance while the second
is delayed. During the direct observation format, supervisors have a rather short and rigid time for feedback and teaching and may feel under pressure, preventing them to discuss and match their observations to students’ needs. The video-based format provides a longer and quieter time during which the supervisor may have more freedom and flexibility on how to organize time between observation, discussion, and teaching.

More importantly, supervisors may be encouraged to rely on different conceptual frameworks of learning according to the type of format. Argyris and Schön described two ways of learning inside an organization (21): a single-loop learning in which errors are detected and reflection is directed towards an immediate solution (problem-solving); a double-loop learning where errors are detected and reflection results more into questioning and even changing the overall framework than finding the more effective strategy. In the direct observation format, the fact that the supervisor spends more time talking and giving advice suggests a single-loop type of learning/teaching. Because the video-based format offers the student the opportunity to self-assess and reflect on his or her own performance, the supervisor may be more likely to act as a facilitator and to question students’ assumptions and frames of reference in a double-loop approach rather than providing immediate solutions. This results in more time for reflection, exchange, and elaboration of concepts and frames.

The reflective practice framework may also bring additional explanation to our observations (22, 23). In direct observation feedback, the student self-assessment and reflection ‘in action’ may be insufficient to provide the supervisor with elaborated materials to discuss. In the video-based feedback session, the students may have had sufficient time and opportunity to reflect on their actions, thus allowing for discussion about more elaborated self-reported materials during the feedback session.

Finally, because the video displays concrete and observable behaviors in a neutral way, it also avoids any source of misunderstanding between the student and the supervisor (9) and may stimulate the sharing of perceptions. This can be especially useful to teach dimensions such as communication and professionalism. Such issues often remain unaddressed during feedback not only because of supervisors’ lack of training or frames of reference (24) but also perhaps because of their reluctance and fear of being threatening or intrusive. This is of importance since several studies have shown a decline in empathy among students and residents during undergraduate and graduate medical education (25). Self-reflection during review of student videos linked to supervisor or peer feedback seems to be an effective and valued way to learn communication and professional attitudes (26, 27) and is more beneficial than traditional feedback on students’ communication skills (9). Video-based stimulated recall has also been shown to be useful to stimulate clinical reasoning or to facilitate development of shared cognition (28, 29). Thus, it seems that video-based feedback, by stimulating self-assessment, improves both self-awareness and skills (9, 30), especially if feedback is coupled to explicit performance expectations and benchmarks (8).

There are several limitations to our study. The sample of supervisors involved and the number of feedback sessions per supervisor analyzed were small. Second, the analysis of the feedback content, focused on counting the elements addressed during feedback, gives only a limited view of what was discussed during the session. Third, we only linked supervisors’ feedback to student perceptions of improvement and not to objective outcomes such as skill improvement or performance change. Finally, the fact that the study was conducted in only one medical school may limit the generalizability of our findings.

Implications for practice

Our findings suggest that these two feedback formats may be applied differently according to the goals of the teaching session. For daily clinical supervision in busy clinics, short, immediate feedback following direct observation focusing on ‘single-loop’ reflection appears to be the most suitable format given time constraints. It is facilitated by the widespread use of the Mini-Clinical Evaluation Exercise (Mini-CEX), a work-based assessment used to evaluate a trainee’s clinical performance in real-life settings (31). Delayed and video-based feedback is more adapted for longer and less frequent sessions addressing issues such as professionalism, communication, and clinical reasoning where prior self-assessment, analysis and discussion of performance, and questioning of frames of reference may both increase students’ self-awareness and facilitate pedagogical diagnosis and remediation.

Conflict of interest and funding

The authors declare that they have no competing interests. The Edmond J. Safra Philanthropic Foundation supported the cost of the data collection.

References