Supplementary material. Appendix 3

Results: SUPPLEMENTARY INFORMATION

Comparison 2: The effect of verbal face-to-face feedback on performance, compared to alternative feedback:

Included studies

Participants

Included studies involved medical students (1076, 55%) in 14 studies,[1-14] mixed health professional students (640, 32%) in one study,[15] pharmacy students (153, 8%) in one study[16] and doctors (105, 5%) in four studies.[17-20]

Participants were novice to the task in 11 studies (11/20, 55%). Three studies documented prior experience: one study involved attending physicians teaching on ward rounds with a range of experience[20] and two studies documented previous training including CPR[12] and history taking and communication skills in medical students.[11] The remaining six studies did not report this information. One of these studies[15] involved teaching CPR to first year health professional students across a mix of disciplines, some of whom may have had prior experience. One study[19] involved evaluating professional and clinical skills in first year paediatric residents who likely had relevant training as medical students. In two of these studies, the baseline performance of junior medical students' attempting surgical knot tying was poor, which suggest limited prior experience.[2, 3] In the last two studies there was no information on prior experience: one assessed a simulated medication consultation by third year pharmacy students[16] and another[9] assessed case presentation skills in third year medical students in their paediatric attachment.

Workplace tasks and Settings

All studies included assessment of a discrete task except two studies which involved longitudinal evaluations.[19, 20] Three studies evaluated performance in a clinical practice setting

(involving teaching skills[20] professional and communication skills[19] and oral case presentations[9]) and the remaining 17 assessed performance in a simulated environment.[1-8, 10-18] Simulated surgical tasks included suturing and/or knot tying,[1-3, 10, 14] bench top surgical procedures such as vascular anastomosis,[18] flexible ureteroscopy for urolithiasis,[5] renal artery angioplasty and stent placment,[17] or surgery using a VR simulator for a laparoscopic salpingectomy.[13] Simulated critical care tasks included basic life support (BLS)/CPR,[12, 15] intubation[6] and pharmacist-patient consultation.[16] The remaining simulated tasks included a hearing test,[7, 8] simulated patient consultation[11] and nasogastric tube insertion.[4]

Interventions

Each study included at least one verbal face-to-face feedback group, in accordance with the inclusion criteria.

Some studies investigated straightforward variations in feedback, including differences in frequency (low or high[4]), stage of practice (early or late[5]), different feedback models ('learning conversation' compared with 'feedback sandwich' frameworks[15]), source expertise (expert or peer[17]) and expert feedback compared to compliments.[2, 3] Another collection of studies explored the effect of adding expert feedback to other interventions, such as in addition to simulator performance data[12, 13] or to written feedback;[19, 20] or adding expert review of a participant's performance video to a practice session in which expert feedback was available.[18] One study[1] compared verbal feedback by an expert who had just directly observed the performance, with written feedback emailed later that day by another expert who watched a video of the performance. Other studies explored more complex phenomena. One study[6] compared two feedback variations in different combinations across four groups. One variation compared an evaluative verbal comment from an expert, to a written numerical performance rating. The second variation involved an individual comparing their performance evaluation to either their own previous attempts (highlighting individual progress) or to expected

performance at student, resident or specialist level (comparison with others). Another research group investigated two complex influences in separate studies. One study[8] examined how the credibility of the feedback provider (high or low credibility) influenced learner outcomes. The other study[7] examined the effect of phrasing corrective information in different ways, so in one intervention corrective information was framed within a positive phrase whereas in the other, it was framed within a critical phrase.

Research funding

One study[12] was loaned a device by Philips as detailed earlier, seven studies received funding from independent institutions,[4, 6, 9, 13, 14, 19, 20] six studies did not receive any funding[2, 5, 7, 8, 11, 17] and six studies did not report information on funding.[1, 3, 10, 15, 16, 18]

Risk of bias

The risk of bias assessment for the comparison of verbal face-to-face feedback to alternative feedback is presented in Figure 3. Seven described an adequate method for randomised sequence generation and allocation concealment, so we rated these studies as 'low risk'.[1, 7-9, 12, 13, 15] Two studies had adequate random sequence generation, which we rated 'low risk' but had insufficient information on allocation concealment, which we rated 'unclear risk'.[11, 19] The remaining studies simply stated participants were 'randomised' and had insufficient information on allocation concealment, both of which we rated 'unclear risk'. Two studies described inconsistencies with randomisation, so these were rated 'high risk' of bias for sequence generation and allocation concealment.[6, 16] There was unequal baseline performance between groups reported in one study[8] and identified from another study's data (obtained from authors).[1] No statistically significant differences in baseline performance between groups were reported in seven studies [2, 4, 5, 8, 14, 19, 20] and baseline performance was not reported in eleven

studies.[3, 6, 9-13, 15-18] None could blind participants or research team members due to the face-to-face feedback interventions. However we thought this was not likely to influence the outcome as implementation and adherence to the intervention were not affected in all studies, which were rated 'low risk', except one in which some participants may not have experienced the intervention they were allocated to, so it was rated 'unclear'.[9] The outcome was assessed by blinded assessors or machines in all studies, which were rated 'low risk' except two studies that did not explicitly describe blinded assessors, which were rated 'unclear'[8, 20] and four studies that seemed to have assessors who were aware of participant allocation, so these were rated 'high risk'.[10, 11, 16] All had high proportions of participant completion data except three[7, 8, 18] and one report provided insufficient information.[3] Three studies had prior published protocols and reported all outcomes as planned, so they were rated 'low risk'.[13, 15, 19] All of the others did not have a prior published protocol but did present outcomes as expected and were rated as 'unclear',[1-9, 11, 12, 14, 16-18, 20] except one study which was rated as 'high risk'.[10]

In summarising the risk of bias across domains within each study, two studies were rated as low risk[13, 15] as all domains were rated as 'low risk' of bias, seven studies were rated as 'high risk' because at least one domain was rated as 'high risk',[6-8, 10, 16, 18] and the remaining studies were rated as 'unclear' as they had at least one domain with 'unclear' risk but no 'high risk' ratings.

References

1. Al-Jundi W, Elsharif M, Anderson M, et al. A Randomized Controlled Trial to Compare e-Feedback Versus "Standard" Face-to-Face Verbal Feedback to Improve the Acquisition of Procedural Skill. *Journal of Surgical Education* 2017;74(3):390-97. doi: https://dx.doi.org/10.1016/j.jsurg.2016.11.011
2. Boehler ML, Rogers DA, Schwind CJ, et al. An investigation of medical student reactions to feedback: a randomised controlled trial. *Medical Education* 2006;40(8):746-9. doi: https://dx.doi.org/10.1111/j.1365-2929.2006.02503.x
3. Rogers DA, Boehler ML, Schwind CJ, et al. Engaging medical students in the feedback process. *American Journal of Surgery* 2012;203(1):21-5. doi: https://dx.doi.org/10.1016/j.amjsurg.2011.07.009

4. Bosse HM, Mohr J, Buss B, et al. The benefit of repetitive skills training and frequency of expert feedback in the early acquisition of procedural skills. *BMC Medical Education* 2015;15:22. doi: https://dx.doi.org/10.1186/s12909-015-0286-5
5. Lee JY, McDougall EM, Lineberry M, et al. Optimizing the Timing of Expert Feedback During Simulation-Based Spaced Practice of Endourologic Skills. *Simulation in Healthcare: The Journal of The Society for Medical Simulation* 2016;11(4):257-63. doi: https://dx.doi.org/10.1097/SIH.0000000000000165
6. Manzone J, Tremblay L, You-Ten KE, et al. Task- versus ego-oriented feedback delivered as numbers or comments during intubation training. *Medical Education* 2014;48(4):430-40. doi: https://dx.doi.org/10.1111/medu.12397
7. Ridder J, Peters C, Stokking K, et al. Framing of feedback impacts student's satisfaction, self-efficacy and performance. *Advances in Health Sciences Education* 2015;20(3):803-16. doi: 10.1007/s10459-014-9567-8
8. van de Ridder J, Berk FC, Stokking KM, et al. Feedback providers' credibility impacts students' satisfaction with feedback and delayed performance. *Medical Teacher* 2015;37(8):767-74. doi: http://dx.doi.org/10.3109/0142159X.2014.970617
9. Sox CM, Dell M, Phillipi CA, et al. Feedback on oral presentations during pediatric clerkships: a randomized controlled trial. *Pediatrics* 2014;134(5):965-71. doi: https://dx.doi.org/10.1542/peds.2014-1209
10. O'Connor A, Schwaitzberg SD, Cao CG. How much feedback is necessary for learning to suture? *Surgical Endoscopy* 2008;22(7):1614-9. doi: https://dx.doi.org/10.1007/s00464-007-9645-6
11. Ozcakar N, Mevsim V, Guldal D, et al. Is the use of videotape recording superior to verbal feedback alone in the teaching of clinical skills? *BMC Public Health* 2009;9:474. doi: https://dx.doi.org/10.1186/1471-2458-9-474
12. Pavo N, Goliasch G, Nierscher FJ, et al. Short structured feedback training is equivalent to a mechanical feedback device in two-rescuer BLS: a randomised simulation study. *Scandinavian Journal of Trauma, Resuscitation & Emergency Medicine* 2016;24:70. doi: https://dx.doi.org/10.1186/s13049-016-0265-9
13. Strandbygaard J, Bjerrum F, Maagaard M, et al. Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: a randomized trial. *Annals of Surgery* 2013;257(5):839-44. doi: https://dx.doi.org/10.1097/SLA.0b013e31827eee6e
14. Xeroulis GJ, Park J, Moulton CA, et al. Teaching suturing and knot-tying skills to medical students: a randomized controlled study comparing computer-based video instruction and (concurrent and summary) expert feedback. *Surgery* 2007;141(4):442-9. doi: https://dx.doi.org/10.1016/j.surg.2006.09.012
15. Baldwin LJ, Jones CM, Hulme J, et al. Use of the learning conversation improves instructor confidence in life support training: An open randomised controlled cross-over trial comparing teaching feedback mechanisms. *Resuscitation* 2015;96:199-207. doi: https://dx.doi.org/10.1016/j.resuscitation.2015.08.010
16. DeLucenay AJ, Conn KM, Corigliano A. An evaluation of the impact of immediate compared to delayed feedback on the development of counselling skills in pharmacy students. *Pharmacy Education* 2017;17(1):322-28.
17. Boyle E, O'Keeffe DA, Naughton PA, et al. The importance of expert feedback during endovascular simulator training. *Journal of Vascular Surgery* 2011;54(1):240-48.e1. doi: https://dx.doi.org/10.1016/j.jvs.2011.01.058
18. Backstein D, Agnidis Z, Sahdu R, et al. Effectiveness of repeated video feedback in the acquisition of a surgical technical skill. *Canadian Journal of Surgery* 2005;48(3):195-200.
19. Brinkman WB, Geraghty SR, Lanphear BP, et al. Effect of multisource feedback on resident communication skills and professionalism: a randomized controlled trial. *Archives of*

Pediatrics & Adolescent Medicine 2007;161(1):44-9. doi: https://dx.doi.org/10.1001/archpedi.161.1.44
20. Skeff KM. Evaluation of a method for improving the teaching performance of attending physicians. *American Journal of Medicine* 1983;75(3):465-70.