Debriefing for technology-enhanced simulation: a systematic review and meta-analysis

Adam Cheng,[1] Walter Eppich,[2] Vincent Grant,[1] Jonathan Sherbino,[3] Benjamin Zendejas[4] & David A Cook[5]

OBJECTIVES Debriefing is a common feature of technology-enhanced simulation (TES) education. However, evidence for its effectiveness remains unclear. We sought to characterise how debriefing is reported in the TES literature, identify debriefing features that are associated with improved outcomes, and evaluate the effectiveness of debriefing when combined with TES.

METHODS We systematically searched databases, including MEDLINE, EMBASE and Scopus, and reviewed previous bibliographies for original comparative studies investigating the use of TES with debriefing in training health care providers. Reviewers, in duplicate, evaluated study quality and abstracted information on instructional design, debriefing and outcomes. Effect sizes (ES) were pooled using random-effects meta-analysis.

RESULTS From 10 903 potentially eligible studies, we identified 177 studies (11 511 learners) that employed debriefing as part of TES. Key characteristics of debriefing (e.g. duration, educator presence and characteristics, content, structure/method, timing, use of video) were usually incompletely reported. A meta-analysis of four studies demonstrated that video-assisted debriefing has negligible and non-significant effects for time skills (ES = 0.10) compared with non-video-assisted debriefing. Meta-analysis demonstrated non-significant effects in favour of expert modelling with short debriefing in comparison with long debriefing (ES range = 0.21–0.74). Among studies comparing terminal with concurrent debriefing, results were variable depending on outcome measures and the context of training (e.g. medical resuscitation versus technical skills). Eight additional studies revealed insight into the roles of other debriefing-related factors (e.g. multimedia debriefing, learner-led debriefing, debriefing duration, content of debriefing). Among studies that compared simulation plus debriefing with no intervention, pooled ESs were favourable for all outcomes (ES range = 0.28–2.16).

CONCLUSIONS Limited evidence suggests that video-assisted debriefing yields outcomes similar to those of non-video-assisted debriefing. Other debriefing design features show mixed or non-significant results. As debriefing characteristics are usually incompletely reported, future debriefing research should describe all the key debriefing characteristics along with their associated descriptors.

Medical Education 2014; 48: 657–666
doi: 10.1111/medu.12432

Discuss ideas arising from the article at www.mededuc.com 'discuss'

[1]Department of Paediatrics, University of Calgary, Calgary, Alberta, Canada
[2]Department of Paediatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA
[3]Department of Medicine, McMaster University, Hamilton, Ontario, Canada
[4]Department of Surgery, Mayo Clinic College of Medicine, Rochester, MN, USA
[5]Department of Internal Medicine, Mayo Clinic College of Medicine, Rochester, MN, USA

Correspondence: Adam Cheng, MD, Alberta Children's Hospital, 2888 Shaganappi Trail NW, Calgary, Alberta T3B 6A8, Canada. Tel: 00 1 403 955 2633; E-mail: chenger@me.com

INTRODUCTION

Debriefing plays an essential role in technology-enhanced simulation (TES) education for health care providers.[1–6] The debriefing provides a forum for learners to reflect on action, identify performance gaps, discuss areas for improvement, and consolidate knowledge and skills so that the latter can be applied in real practice to improve health care and patient outcomes.[7–10] Although debriefing is often considered to represent a subgroup of feedback techniques,[1,2,11] this article will treat these as distinct instructional activities. We propose that the hallmark of debriefing is the interactive, bi-directional and reflective nature of discussion,[3–5,7–10,12] whereas feedback without debriefing is a unidirectional communication about the recipient's behaviour.[13,14]

Recent reviews of the simulation health care literature have highlighted the benefits of feedback as part of TES education.[12,15] These studies included analysis that contained elements of debriefing, but did not look specifically at the debriefing activities. Other articles have attempted to clarify how debriefing might be implemented for different contexts, learners and objectives.[3–6] These narrative reviews describe various methods of debriefing and the theoretical underpinning of these methods. However, none of these articles quantitatively synthesise the results of existing debriefing research, nor do they describe the extent and quality of reported research. A clear picture of how debriefing has been used in TES for health care providers and the factors that make debriefing effective is essential in order to elucidate how debriefing can be improved, as well as to inform future research in this area.

To address this need, we conducted a systematic review with the goals of: (i) describing the characteristics of debriefing in studies involving TES with health care providers; (ii) identifying debriefing characteristics associated with improved outcomes, and (iii) evaluating the effectiveness of TES with debriefing for teaching health care professionals in comparison with no intervention or non-simulation instruction.

METHODS

This review drew from studies identified in a comprehensive systematic review of TES,[15] supplemented by a targeted search for additional relevant studies. This review was planned, conducted and reported in adherence with PRISMA (*p*referred *r*eporting *i*tems for *s*ystematic reviews and *m*eta-*a*nalyses) standards of quality for reporting meta-analyses.[16]

Study eligibility

We included comparative studies published in any language if they investigated the use of debriefing in conjunction with TES training. Many definitions of debriefing exist. The common themes refer to 'a facilitated or guided reflection in the cycle of experiential learning',[3,7] 'that helps learners develop and integrate insights…into later action',[7] and 'helps participants understand, analyse and synthesise what they thought, felt and did'[7,8] through facilitated discussion between two or more individuals.[3–5,7–10] For the purposes of this article, we define debriefing as a discussion between two or more individuals in which aspects of a performance are explored and analysed with the aim of gaining insights that impact the quality of future clinical practice. We define TES as the application of an educational tool or device with which the learner physically interacts to mimic an aspect of clinical care.[15] We included studies using single group pre-test/post-test, two-group non-randomised, and randomised designs.

Study identification

We employed two separate search strategies to identify studies of interest. The first search, which used an end date of May 2011,[15] sought studies of TES broadly, and included terms such as 'simulation', 'training' and 'manikin' among others. From this search, we selected studies relevant to our research question. To identify recently published articles focused on evaluating different approaches to debriefing, our second search queried MEDLINE using the search terms 'simulat*' AND 'debrief*' and sought studies published between 1 January 2011 and 15 December 2012. We also examined the entire reference lists of several published reviews of debriefing.[3–5,10]

Study selection

Two independent reviewers screened for inclusion the titles and abstracts of all potentially eligible studies, referring to the full text when determinations of inclusion were uncertain. Conflicts were resolved by discussion and by reaching consensus. The chance-adjusted inter-rater agreement for inclusion in the focused search, determined using the intraclass correlation coefficient (ICC), was 0.91.

Data extraction

We extracted information from each study independently and in duplicate; all conflicts were resolved by reaching consensus. We abstracted information on the clinical topic, training level of learners, instructional design features of simulation training, study design, method of group assignment, outcomes type and methodological quality.[15] Methodological quality was graded using the Medical Education Research Study Quality Instrument (MERSQI)[17] and an adaptation of the Newcastle–Ottawa Scale (NOS)[18] for cohort studies. Data were abstracted separately for learning outcomes of knowledge, skills, behaviours with patients, and direct effects on patients. Skills outcomes were further classified according to whether they related to time (e.g. time to complete the task), process (e.g. performance rating in a simulated setting) or product (e.g. successful task completion or major errors) outcomes, and behaviours with real patients were similarly classified as time- or process-related (e.g. in the real patient setting).[15] To code characteristics of debriefing, we used as a starting point the features defined in a previous review article by Raemer *et al.*[4] which breaks down the debriefing into the categories of: who (*debriefer*); what (*methods/content*); when (*timing*); where (*environment*), and why (*theory*). We built on this work by adding two new categories (*debriefing duration* and *educator presence*), and separating debriefing method and content into distinct categories. Additionally, we suggest descriptors to report within each category. All characteristics of debriefing were extracted independently from each study, and all conflicts were resolved by discussion and consensus. We coded the following features of the debriefing: duration (ICC: 0.90); use of video (ICC: 0.73); educator presence (ICC: 0.58); educator characteristics (ICC: 0.77); debriefing content (ICC: 0.86); debriefing method (ICC: 0.48), and timing of debriefing (ICC: 0.99).

Data synthesis

A standardised mean difference (Hedges' *g* effect size) was calculated for each comparison.[19,20] We used a random-effects meta-analysis to quantitatively pool results and organised these analyses by comparison (i.e. with no intervention or with another form of instruction) and outcome. For the studies that made comparisons with no intervention, we conducted subgroup analyses based on key study design and instructional design features. We also performed sensitivity analyses excluding the results of studies with imprecise effect size (ES) calculations.[15] We did not perform subgroup analyses for studies with active comparison because we found only a few eligible articles per analysis. For studies that compared different methods or types of debriefing in TES, we identified themes of comparison, such as video versus non-video debriefing, and then pooled the results of related studies using meta-analysis. Studies that were not conceptually aligned or which seemed to have context-specific drivers were summarised in a narrative synthesis.

We quantified between-study inconsistency for analyses of three or more studies using the I^2 statistic,[21] estimating the percentage of variability not attributable to chance. I^2 values of > 50% indicate large inconsistency or heterogeneity. We used SAS Version 9.1.3 (SAS Institute, Inc., Cary, NC, USA) for all analyses. Statistical significance was defined by a two-sided α-value of 0.05. Interpretations of clinical significance emphasized confidence intervals (CIs) in relation to Cohen's ES classifications (> 0.8, large; 0.5–0.8, moderate; 0.2–0.5, small; < 0.2, negligible).[22]

RESULTS

Trial flow

We identified 177 studies that employed debriefing as part of TES, reflecting data from 11 511 learners (Fig. S1). Of these, 108 (61%) studies compared TES and debriefing with no intervention, 27 (15%) made comparisons with another instructional modality, and 48 (27%) made comparisons with another form of TES (some studies made multiple types of comparisons). Table S1 summarises the key features of these studies and Appendix S1 provides a complete listing of references.

Study characteristics: included studies

The most common types of learners were nurses or nursing students (n = 3757), postgraduate physician trainees (n = 2990), and medical students (n = 2428). Sixty-eight studies (38%) focused on resuscitation training, 38 (21%) focused on communication/team skills and 20 (11%) focused on anaesthesia (Table S1). Only 11 studies (6%) focused on surgical skills or endoscopy. A total of 250 distinct outcomes were reported, including 125 process skills and 20 outcomes assessed in the context of actual patient care.

Debriefing characteristics

Debriefing characteristics were incompletely reported in most studies; 116 studies (66%) reported four or fewer of the seven different debriefing characteristics (Table 1). Seventy-one (40%) studies reported the use of video playback during debriefing, 41 (23%) described the duration of debriefing, 68 (38%) described the background characteristics of the educator and 50 (28%) reported the structure or method of debriefing. The duration of debriefing and method of debriefing were the least frequently reported characteristics. Only five studies (3%) reported all seven debriefing characteristics; three of these were studies that specifically altered one characteristic of debriefing in order to determine its effect on learning.[23–25] The other two studies involved simulated cardiopulmonary arrest: one used video-assisted debriefing for nursing students[26] and the other utilised 'debriefing with good judgement' in groups of medical students.[27]

Study quality

Data on the methodological quality of the included studies are summarised in Table S2. The mean ± standard deviation (SD) NOS (maximum 6 points) and MERSQI (maximum 18 points) scores for study quality were 2.7 ± 1.7 and 12.0 ± 2.1, respectively.

Synthesis: comparison of simulation and debriefing versus other types of TES

Forty-eight studies (3039 participants) compared TES with debriefing with other types of simulation-based instructional modalities. We focused on studies that assessed different types of debriefing, and conducted meta-analyses pooling the results of studies that addressed two common comparison themes.[23,24,28–32]

Comparison of video versus non-video debriefing

Four studies compared the use of video-assisted debriefing with non-video-assisted debriefing for TES (Fig. 1),[28–31] and showed minimal difference in benefit between the two methods of debriefing. A meta-analysis of these studies showed negligible and non-significant effects for time skills (ES = 0.10, 95% CI −0.20 to 0.40; p = 0.53; n = 2 studies), and process skills (ES = −0.06, 95% CI −0.28 to 0.17; p = 0.63; n = 4 studies). Confidence intervals excluded effects of moderate magnitude, and results were notably consistent across the four process outcomes (low inconsistency, I^2 = 2%).

Comparison of long debriefing with short debriefing and expert modelling

Three separate studies compared simulation-based instruction plus a long debriefing (> 30 minutes) with short debriefing (< 15 minutes) combined with viewing of expert-modelled performance (Fig. 1).[23,24,32] These studies suggested a possible benefit from expert modelling with short debriefing, although the results were usually not statistically significant. One study found a statistically significant benefit in favour of expert modelling plus short debriefing for knowledge (ES = 0.91, 95% CI 0.11–1.70; p = 0.03). Meta-analyses of other outcomes revealed non-significant effects in favour of expert modelling plus short debriefing, including moderate effects for process skills (ES = 0.74, 95% CI −0.58 to 2.05; p = 0.27; n = 3 studies), and

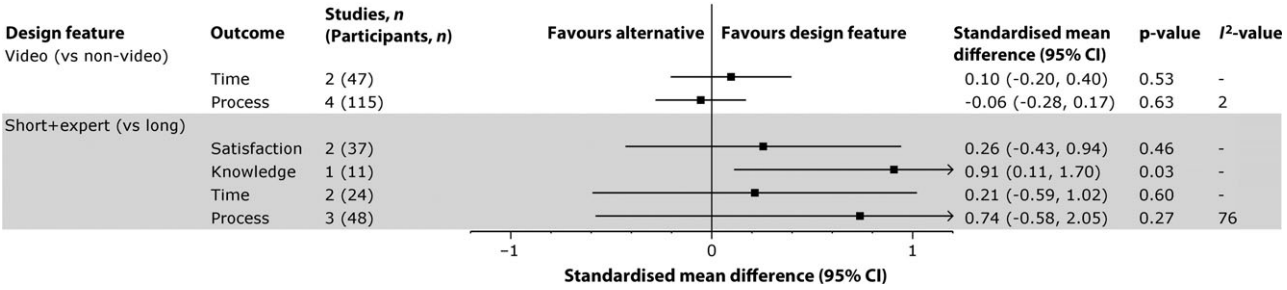

Figure 1 Outcomes of studies comparing alternative approaches to simulation-based debriefing. The figure summarises the results of seven studies, each focused on evaluating one of two key design features using one or more outcomes. The point estimates and 95% confidence intervals (CIs) reflect meta-analytic pooling when they refer to more than one study. Design themes are: (i) video versus non-video debriefing (a positive number favours video debriefing), and (ii) short debriefing (< 15 min) with expert modelling versus long debriefing (> 30 min) (a positive number favours short debriefing with expert modelling)

small effects for participant satisfaction (ES = 0.26, 95% CI −0.43 to 0.94; p = 0.46; *n* = 2 studies) and time outcomes (ES = 0.21, 95% CI −0.59 to 1.02; p = 0.6; *n* = 2 studies).

Other debriefing themes

Several additional studies on debriefing that were not included in the meta-analysis revealed further insight into factors that may impact the effectiveness of debriefing in TES.[25,33–41]

Terminal versus concurrent debriefing

Three studies compared the use of terminal with concurrent debriefing for TES,[33–35] and showed variable results depending on the context and outcome measure. Terminal debriefing was more effective than concurrent debriefing in the context of medical resuscitation[33] (ES = 0.34, satisfaction), whereas concurrent debriefing was more effective for endoscopy skills (ES = −0.51, time skills; ES = −0.26, process skills).[35] A different study reported no difference for process skills in the context of suturing and knot-tying (ES = 0).[34]

Personalised oral versus multimedia instruction

Two studies compared the use of a personalised oral debriefing with an educator after simulation with multimedia instruction with no educator present.[34,36] Personalised oral debriefing was favoured over multimedia use with no educator in the context of suturing and knot-tying skills and anaesthesia resuscitation (ES range = 0.13–0.16, process skills).[34,36]

Self- versus instructor-led debriefing
A comparison of self- versus instructor-led debriefing favoured instructor-led debriefing for process skills (ES = 0.18).[25]

In-person versus web-conferenced instructors
Debriefing conducted by an instructor via web conference was more effective than in-person debriefing (ES = 0.96, process skill).[37]

Audiovisual feedback during simulation plus debriefing
The addition of concurrent audiovisual feedback (during simulation) for chest compressions followed by debriefing enhanced the effectiveness of learning compared with the provision of no audiovisual feedback during simulation (ES = 0.66, process skill).[38]

Amount of debriefing
Limited feedback in the context of laparoscopic skills training was found to be superior to an interactive debriefing session, leading to quicker achievement of proficiency, and decreasing the costs of laparoscopic skills training (ES = 0.24).[39]

Debriefing style and content
Learners favoured debriefing with a greater emphasis on reflection (ES = 0.27, knowledge) over debriefing focused on performance critique,[40] although a different study showed a greater learner preference for debriefings that were more technical in nature over cognitive debriefings (ES = 0.55, satisfaction).[41]

Synthesis: comparison of simulation and debriefing versus no intervention

Figure S2 and Tables S3 and S4 summarise the meta-analyses of 108 studies comparing TES plus debriefing with no intervention for outcomes of knowledge and process skills. In summary, simulation training with debriefing was associated with favourable and statistically significant effects for nearly all outcomes, although there was substantial inconsistency between studies (I^2 range: 70–91%). Specifically, results by outcome were: knowledge, ES = 0.86 (95% CI 0.66–1.06; p < 0.001; *n* = 32 studies); process skill, ES = 1.01 (95% CI 0.88–1.14; p < 0.001; *n* = 80 studies); time skills, ES = 0.75 (95% CI 0.39–1.11; p < 0.001; *n* = 11 studies); product skills, ES = 2.16 (95% CI 0.72–3.60; p = 0.003; *n* = 2 studies); behaviour process, ES = 0.56 (95% CI 0.31–0.80; p < 0.001; *n* = 9 studies); behaviour time, ES = 0.28 (95% CI −0.33 to 0.89; p = 0.37; *n* = 2 studies), and patient effects, ES = 0.30 (95% CI 0.04–0.55; p = 0.02; *n* = 7 studies). Subgroup analyses exploring the large degree of inconsistency revealed no significant interaction with research method (Tables S3 and S4).

Synthesis: comparison of simulation and debriefing versus other forms of instruction

Twenty-seven studies compared TES plus debriefing with non-simulation instructional modalities (Table S5). Results showed small effects in favour of TES with debriefing for knowledge, time and process outcomes, and moderate effects for satisfaction.

DISCUSSION

We sought to systematically review and synthesise the evidence regarding the use of debriefing in TES. The characteristics of debriefing were inconsistently reported; only a small minority of studies described all seven debriefing characteristics. Our analysis of studies comparing different forms of debriefing for TES suggest that video-assisted debriefing yields outcomes similar to those of non-video-assisted debriefing. Other debriefing design features show mixed or non-significant results. In comparison with no intervention and other forms of instruction, TES with debriefing was associated with effects that were uniformly favourable, but somewhat variable in magnitude.

Limitations and strengths

Many other definitions of debriefing exist, and we might have included or excluded studies differently had we used another definition. The results of our meta-analysis examining different types of debriefing methods should be interpreted with the knowledge that few studies were included in each analysis. Additionally, debriefing characteristics were poorly described in many studies, which made it difficult at times to identify these features with high reliability. Further work in this area will help to shed light on the relative contributions of the various characteristics of debriefing.

We found substantial inconsistency between studies in most analyses, which is likely to stem from the differences in instructional and study designs, and types of learners enrolled. Although this is a liability when interpreting the precise magnitude of effect, it is also a strength insofar as it allows our results to generalise broadly. Moreover, for studies that compared outcomes with those of no intervention, our findings suggest inconsistency in the magnitude but not in the direction of effect. We elected not to evaluate for publication bias because these methods are unreliable in the presence of high inconsistency among studies.[42]

The strengths of our study include its exhaustive initial literature search, its focused search for recent high-impact studies, the duplicate review at all stages, and the detailed quantitative synthesis of findings.

Integration with prior work

Several narrative reviews have described the various methods of debriefing,[3–6,12] and other authors have proposed specific methods of debriefing[7–10,12] supported by evidence from other fields or personal experience. One other study used qualitative methods to describe the pattern of interactions between instructors and participants.[10] No previous articles on debriefing have enumerated the use of debriefing in simulation-based research, described how debriefing characteristics are reported in existing literature, or attempted to quantitatively synthesise the relative benefits of various types of debriefing characteristics. As such, our current work represents an important contribution to the field and may help to shape the nature of future education and research.

Previous systematic[1,12] and non-systematic[2–4] reviews suggest that feedback, which in some cases included debriefing, is a critical feature of TES that contributes to favourable learning outcomes. Our study confirms this notion and builds on these prior efforts by focusing specifically on studies with debriefing (and excluding studies of unidirectional feedback). The quality of research reported in our study parallels that of research cited in previous systematic reviews of the TES literature.[14,15] Previous reviews using the MERSQI have reported average scores ranging from 9.6 to 12.3,[14,15,17,43,44] which are comparable with the mean MERSQI score of 12 in the present study.

Implications for current practice and future research

We see several implications for the field of debriefing related to TES. Firstly, although video debriefing has increased in popularity,[28–31] our analyses (which demonstrate high consistency between studies) suggest that video debriefing has negligible benefit in comparison with non-video debriefing. Given the limited number of studies, it is possible that these findings are specific to the learner group, topic or method of video use described in these studies. As such, further study to determine specific content, learner groups or environments that might benefit from video debriefing would be beneficial. Specific methods for using video debriefing (e.g. when video should be used, how discussion should be structured around video) require exploration to determine the optimal means of improving learning outcomes. Pending further research, simulation educators should carefully weigh the benefits of video debriefing against the costs of video hardware and software before making the decision to integrate video debriefing into simulation-based learning.

Table 1 Description of debriefing characteristics

Debriefing characteristic/ Response	All studies with this feature (n = *177*), n *(%)*	Studies using video debriefing with this feature, (n = *71*), n *(%)**
Duration of debriefing		
Not described	136 (77%)	53 (75%)
Short (≤ 15 minutes)	15 (8%)	2 (3%)
Long (> 15 minutes)	21 (12%)	12 (17%)
Both (in alternate interventions)†	5 (3%)	4 (6%)
Educator presence		
Not described	23 (13%)	10 (14%)
Educator present	151 (85%)	60 (85%)
Educator not present	1 (1%)	0
Both (in alternate interventions)†	2 (1%)	1 (1%)
Educator characteristics		
Not described	109 (62%)	40 (56%)
Clinical background (e.g. profession or specialty)	40 (23%)	16 (23%)
Debriefing experience (e.g. specific training in debriefing, number of debriefing sessions facilitated)	16 (9%)	9 (13%)
Both clinical and debriefing experience	8 (5%)	5 (7%)
Other (e.g. advanced degree)	3 (2%)	1 (1%)
Both (in alternate interventions)†	1 (1%)	0
Content of debriefing		
Not described or vague	34 (19%)	13 (18%)
Medical management or clinical decision making	78 (44%)	22 (31%)
CRM (e.g. teamwork, leadership)	24 (14%)	16 (23%)
Both medical management and CRM	32 (18%)	18 (25%)
Other content (e.g. technical skills)	7 (4%)	1 (1%)
Both (in alternate interventions)†	2 (1%)	1 (1%)
Structure and method of debriefing		
Not described or vague	127 (72%)	49 (69%)
Advocacy–inquiry ('debriefing with good judgement'[8,9])	4 (2%)	1 (1%)
Plus-Delta[3]	10 (6%)	4 (6%)
Other (e.g. assessment tool-guided, chronologically guided)	34 (19%)	17 (24%)
Both (in alternate interventions)†	2 (1%)	0
Timing of debriefing		
Not described	13 (7%)	2 (3%)
Terminal (after simulation)	152 (86%)	69 (97%)
Concurrent (simulation interrupted with debriefing in or during simulation)	6 (3%)	0
Both terminal and concurrent	3 (2%)	0
Both (in alternate interventions)†	3 (2%)	0

CRM = crisis resource management
* Five studies employed both video and non-video debriefing
† This debriefing characteristic varied between study arms in these studies

Secondly, a group of three studies showed promising (but somewhat inconsistent) effects when combining short debriefing with instructors' expert modelling of clinically appropriate behaviours in selected clinical vignettes.[23,24,32] In these studies, the expert modelling was followed by a modified debriefing in which students were allowed to ask questions related to the actions and thought processes of the instructor team. This research suggests a potential role for expert modelling as an adjunct to debriefing, and draws attention to the benefit of blending other learning modalities with debriefing. Future research could further explore the optimal context for other instructional methods to augment debriefing.

Thirdly, several studies have attempted to delineate the ideal timing of debriefing.[33–35] The variability in the results of these studies suggests that the ideal timing is likely to depend on various factors, including the topic of learning (e.g. technical skills versus behavioural skills), individual versus team-based learning, and task complexity. Individual learners asked to practise more basic tasks (e.g. intravenous insertion) may be more suited to the provision of concurrent feedback, whereas a team of health care professionals learning how to manage a complex medical resuscitation may be more suited to terminal debriefing, where an appropriate amount of time can be spent on reflective discussion.[45,46] Educators should consider these issues when deciding between terminal and concurrent debriefing.

Fourthly, debriefing is most commonly used for resuscitation, communication and team skills, and in anaesthesia-based simulations. Conversely, the use of debriefing is infrequent in training for procedural or surgical skills simulations. Whether this divergence is appropriate (e.g. perhaps debriefing is not required for procedural skills) or not might be the topic of future research to define which clinical topics are best suited to debriefing, and the relative benefits of interactive debriefing and unidirectional feedback, respectively.

Our study identified that debriefing characteristics are inconsistently reported when used in TES. We build on Raemer *et al.*'s[4] previous categorization of debriefing characteristics by adding additional categories and suggesting specific descriptors for all of the categories. We suggest that authors should consistently use these key descriptors in future research (Table 1). A standardised method of reporting will help to advance the field by facilitating better understanding of how debriefings are conducted, thus allowing educators to closely replicate debriefing methods that are proven to be effective.

Lastly, our literature review identified that only 10% of all simulation studies involving debriefing compared one method or style of debriefing with another. In addition to the themes noted above, the field would benefit from evidence exploring other themes, such as the effective implementation of learner-centred debriefing or scripted debriefing, the relative merits of different debriefing methods, and how the context and topic of instruction influence the need for and effectiveness of various debriefing approaches.

Contributors: AC contributed to the conceptualising and design of the study, and to data collection and interpretation. WE, VG and JS contributed to the design of the study, and to data collection. BZ contributed to data analysis and interpretation. DAC contributed to the conceptualising and design of the study and to data collection, analysis and interpretation. All authors contributed to the drafting and revision of the paper and approved the final manuscript for publication.
Acknowledgements: the authors would like to acknowledge and thank Ryan Brydges, Stan Hamstra, Rose Hatala, Jason Szostek, Amy Wang and Pat Erwin for their contributions to initial study selection and data abstraction.
Funding: this study was supported by internal funding from the Mayo Clinic Division of General Internal Medicine.
Conflicts of interest: none.
Ethical approval: not applicable.

REFERENCES

1 Issenberg SB, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 2005;**27**:10–28.
2 McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003–2009. *Med Educ* 2010;**44**:50–63.
3 Fanning RM, Gaba DM. The role of debriefing in simulation-based learning. *Simul Healthc* 2007;**2**:115–25.
4 Raemer D, Anderson M, Cheng A, Fanning R, Nadkarni V, Savoldelli G. Research regarding debriefing as part of the learning process. *Simul Healthc* 2011;**6** (Suppl):52–7.
5 Arafeh JM, Hansen SS, Nichols A. Debriefing in simulated-based learning: facilitating a reflective discussion. *J Perinat Neonat Nurs* 2010;**24**:302–9.

6 Lederman LC. Debriefing: toward a systematic assessment of theory and practice. *Simul Gaming* 1992;**23**:145–60.
7 Rudolph JW, Simon R, Raemer D, Eppich WJ. Debriefing as formative assessment: closing performance gaps in medical education. *Acad Emerg Med* 2008;**15**:1–7.
8 Rudolph JW, Simon R, Rivard P, Dufresne RL, Raemer DB. Debriefing with good judgement: combining rigorous feedback with genuine inquiry. *Anesth Clin* 2007;**25**:361–76.
9 Rudolph JW, Simon R, Dufresne RL, Raemer DB. There's no such thing as a 'non-judgemental' debriefing: a theory and method for debriefing with good judgement. *Simul Healthc* 2006;**1**:49–55.
10 Dieckmann P, Molin Friss S, Lippert A, Ostergaard D. The art and science of debriefing in simulation: ideal and practice. *Med Teach* 2009;**31**:e287–94.
11 Chiniara G, Cole G, Brisbin K, Huffman D, Cragg B, Lamacchia M, Norman D. Simulation in healthcare: a taxonomy and a conceptual framework for instructional design and media selection. *Med Teach* 2013;**35**:e1380–95.
12 Dreifuerst KT. The essentials of debriefing in simulation learning: a concept analysis. *Nurs Educ Perspect* 2009;**30**:109–14.
13 Ilgen DR, Fisher CD, Taylor MS. Consequences of individual feedback on behaviour in organisations. *J Appl Psychol* 1979;**64**:349–71.
14 Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek J, Wang AT, Erwin PJ, Hatala R. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. *Med Teach* 2013;**35**:e867–98.
15 Cook DA, Hatala R, Brydges R, Zendejas B, Szostek J, Wang AT, Erwin PJ, Hamstra SJ. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 2011;**3306**:978–88.
16 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analysis: the PRISMA statement. *J Clin Epidemiol* 2009;**62**:1006–12.
17 Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. *JAMA* 2007;**298**:1002–9.
18 Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. *JAMA* 2008;**300**:1181–96.
19 Borenstein M. Effect sizes for continuous data. In: Cooper H, Hedges LV, Valentine JC, eds. *The Handbook of Research Synthesis*, 2nd edn. New York, NY: Russell Sage Foundation 2009;221–35.
20 Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods* 2002;**7**:105–25.
21 Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions 5.1.0. 2011. http://www.cochrane.org/resources/handbook/index.htm. [Accessed 28 February 2013.]
22 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates 1988; 273–406.
23 LeFlore JL, Anderson M, Michael JL, Engle WD, Anderson J. Comparison of self-directed learning versus instructor-modelled learning during a simulated clinical experience. *Simul Healthc* 2007;**2**:170–7.
24 LeFlore JL, Anderson M. Alternative educational models for interdisciplinary student teams. *Simul Healthc* 2009;**4**:135–42.
25 Boet S, Bould D, Bruppacher HR, Desjardins F, Chandra DB, Naik VN. Looking in the mirror: self-debriefing versus instructor debriefing for simulated crises. *Crit Care Med* 2011;**39**:1377–81.
26 Leighton K, Scholl K. Simulated codes: understanding the response of undergraduate nursing students. *Clin Simul Nurs* 2009;**5**:e187–94.
27 DeMaria S Jr, Bryson EO, Mooney TJ, Silverstein JH, Reich DL, Bodian C, Levine AI. Adding emotional stressors to training in simulated cardiopulmonary arrest enhances participant performance. *Med Educ* 2010;**44**:1006–15.
28 Savoldelli GL, Naik VN, Park J, Joo HS, Chow R, Hamstra SJ. Value of debriefing during simulated crisis management: oral versus video-assisted oral feedback. *Anesthesiology* 2006;**105**:279–85.
29 Grant JS, Moss J, Epps C, Watts P. Using video-facilitated feedback to improve student performance following high-fidelity simulation. *Clin Simul Nurs* 2010;**6**:e177–84.
30 Byrne AJ, Sellen AJ, Jones JG, Aitkenhead AR, Hussain S, Gilder F, Smith HL, Ribes P. Effect of videotape feedback on anaesthetists' performance while managing simulated anaesthetic crises: a multicentre study. *Anaesthesia* 2002;**57**:176–9.
31 Sawyer T, Sierocka-Castaneda A, Chan D, Berg B, Lustik M, Thompson M. The effectiveness of video-assisted debriefing versus oral debriefing alone at improving neonatal resuscitation performance: a randomised trial. *Simul Healthc* 2012;**7**:213–21.
32 LeFlore JL, Anderson M. Effectiveness of two methods to teach and evaluate new content to neonatal transport personnel using high-fidelity simulation. *J Perinat Neonat Nurs* 2008;**22**:319–28.
33 Van Heukelom JN, Begaz T, Treat R. Comparison of post-simulation debriefing versus in-simulation debriefing in medical simulation. *Simul Healthc* 2010;**5**:91–7.
34 Xeroulis GJ, Park J, Moulton CA, Reznick RK, LeBlanc V, Dubrowski A. Teaching suturing and knot-tying skills to medical students: a randomised controlled study comparing computer-based video instruction and (concurrent and summary) expert feedback. *Surgery* 2007;**141**:442–9.

35 Walsh CM, Ling SC, Wang CS, Carnahan H. Concurrent versus terminal feedback: it may be better to wait. *Acad Med* 2009;**84** (Suppl):54–7.
36 Welke TM, LeBlanc V, Savoldelli GL, Joo HS, Chandra DB, Crabtree NA, Naik VN. Personalised oral debriefing versus standardised multimedia instruction after patient crisis simulation. *Anesth Analg* 2009;**109**:183–9.
37 Hayden EM, Navedo DD, Gordon JA. Web-conferenced simulation sessions: a satisfaction survey of clinical simulation encounters via remote supervision. *Telemed J E Health* 2012;**18**:525–9.
38 Dine CJ, Gersh RE, Leary M, Riegel BJ, Bellini LM, Abella BS. Improving cardiopulmonary resuscitation quality and resuscitation training by combining audiovisual feedback and debriefing. *Crit Care Med* 2008;**36**:2817–22.
39 Stefanidis D, Korndorffer JR, Heniford T, Scott DJ. Limited feedback and video tutorials optimise learning and resource utilisation during laparoscopic simulator training. *Surgery* 2007;**142**:202–6.
40 Dreifuerst KT. Using debriefing for meaningful learning to foster development of clinical reasoning in simulation. *J Nurs Educ* 2012;**51**:326–33.
41 Bond WF, Deitrick LM, Eberhardt M, Barr GC, Kane BG, Worrilow CC, Arnold DC, Croskerry P. Cognitive versus technical debriefing after simulation training. *Acad Emerg Med* 2006;**13**:276–83.
42 Montori VM, Smieja M, Guyatt GH. Publication bias: a brief review for clinicians. *Mayo Clin Proc* 2000;**75**: 1284–8.
43 Reed DA, Beckman TJ, Wright SM, Levine RB, Kern DE, Cook DA. Predictive validity evidence for medical education research study quality instrument scores: quality of submissions to *JGIM*'s Medical Education Special Issue. *J Gen Intern Med* 2008;**23**:903–7.
44 McKinney J, Cook DA, Wood D, Hatala R. Simulation-based training for cardiac auscultation skills: systematic review and meta-analysis. *J Gen Intern Med* 2013;**28**:283–91.
45 Chang JY, Chang GL, Chien CJC, Chung KC, Hsu AT. Effectiveness of two forms of feedback on training of a joint mobilisation skill by using a joint translation simulator. *Phys Ther* 2007;**87** (4):418–30.
46 Wierinck E, Puttemans V, van Steenberghe D. Effect of reducing frequency of augmented feedback on manual dexterity training and its retention. *J Dent* 2006;**34** (9):641–7.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1. List of all included studies.
Figure S1. Study flow diagram.
Figure S2. Random-effects meta-analysis of studies comparing simulation training plus debriefing with no intervention: knowledge, time, process skill, products skill, learner behaviour and patient outcomes.
Table S1. Description of included studies.
Table S2. Quality of included studies.
Table S3. Random-effects meta-analysis of simulation training with debriefing: knowledge outcomes.
Table S4. Random-effects meta-analysis of simulation training with debriefing: process skill outcomes.
Table S5. Random-effects meta-analysis of studies comparing simulation-based training plus debriefing with non-simulation instruction: satisfaction, knowledge, time, process skill and product skill outcomes.

Received 9 October 2013; editorial comments to author 20 December 2013; accepted for publication 22 January 2014