

INNOVATION

Open Access



# Immediate faculty feedback using debriefing timing data and conversational diagrams

Andrew Coggins<sup>1\*</sup> , Sun Song Hong<sup>1</sup>, Kaushik Baliga<sup>2</sup> and Louis P. Halamek<sup>3</sup>

## Abstract

**Background:** Debriefing is an essential skill for simulation educators and feedback for debriefers is recognised as important in progression to mastery. Existing assessment tools, such as the Debriefing Assessment for Simulation in Healthcare (DASH), may assist in rating performance but their utility is limited by subjectivity and complexity. Use of quantitative data measurements for feedback has been shown to improve performance of clinicians but has not been studied as a focus for debriefer feedback.

**Methods:** A multi-centre sample of interdisciplinary debriefings was observed. Total debriefing time, length of individual contributions and demographics were recorded. DASH scores from simulation participants, debriefers and supervising faculty were collected after each event. Conversational diagrams were drawn in real-time by supervising faculty using an approach described by Dieckmann. For each debriefing, the data points listed above were compiled on a single page and then used as a focus for feedback to the debriefer.

**Results:** Twelve debriefings were included ( $\mu = 6.5$  simulation participants per event). Debriefers receiving feedback from supervising faculty were physicians or nurses with a range of experience ( $n = 7$ ). In 9/12 cases the ratio of debriefer to simulation participant contribution length was  $\geq 1:1$ . The diagrams for these debriefings typically resembled a fan-shape. Debriefings ( $n = 3$ ) with a ratio  $< 1:1$  received higher DASH ratings compared with the  $\geq 1:1$  group ( $p = 0.038$ ). These debriefings generated star-shaped diagrams. Debriefers self-rated DASH scores ( $\mu = 5.08/7.0$ ) were lower than simulation participant scores ( $\mu = 6.50/7.0$ ). The differences reached statistical significance for all 6 DASH elements. Debriefers evaluated the 'usefulness' of feedback and rated it 'highly' ( $\mu = 4.6/5$ ).

**Conclusion:** Basic quantitative data measures collected during debriefings may represent a useful focus for immediate debriefer feedback in a healthcare simulation setting.

**Keywords:** Interprofessional collaborative practice, Debriefing, Faculty development (simulation educator or technician)

\* Correspondence: [andrew.coggins@health.nsw.gov.au](mailto:andrew.coggins@health.nsw.gov.au)

<sup>1</sup>Department of Emergency Medicine, Westmead Hospital, Sydney, NSW 2145, Australia

Full list of author information is available at the end of the article



© The Author(s). 2022 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Providing adult learners with meaningful feedback is likely to be an important contributor to improved future performance [1–3]. Debriefing following simulation-based medical education (SBME) events is a key step in allowing participants to identify performance gaps and sustain good practice [3–5]. To achieve this goal, it is acknowledged that effective debriefing is important [6, 7]. Yet, as is often observed a gap may exist between ideal approaches to debriefing and actual performance [4].

To bridge this gap, a number of debriefing assessment tools provides a guide for rating and reviewing performance [8–10]. The tools available include the Objective Structured Assessment of Debriefing (OSAD) and the Debriefing Assessment for Simulation in Healthcare (DASH). OSAD and DASH assess debriefers on a Likert scale based on a set of ideal behaviours [8, 10]. As a result, they are useful for illustrating concepts to novices and providing a shared mental model of what a good debriefing looks like. However, they are not easily integrated into debriefer feedback, mentoring or coaching [11]. While these tools appear to be widely adopted in the training of debriefers, validation studies were limited to analysis of delayed reviews of recorded debriefings [8–10, 12]. In addition, while the tools may identify areas for the debriefer to improve, the arbitrary scores provided do not necessarily translate to improved future performance. In this study we seek to close this gap by exploring the use of quantitative data measures as a supplementary tool for debriefer feedback.

Current faculty development programmes often use the tools listed above as an aid to achieve improved debriefings [11]. In many programmes, feedback to new debriefers follows direct observation (or video review) by more experienced colleagues. Mentoring may also be useful if provided in a structured manner to help progress new debriefers towards mastery [13]. Coaching using a supportive and pre-agreed approach may also be important for facilitating stepwise improvements in debriefer performance [6, 13, 14]. These 3 strategies (i.e. feedback, mentoring and coaching) are attractive concepts but the best approaches to debriefer faculty development remain uncertain.

Based on an observation of debriefings conducted in various non-healthcare settings, we hypothesised that the use of quantitative data for feedback may provide an additional option for debriefer faculty development [15]. Notably, the current debriefing literature does not extensively report on using such quantitative data for debriefer feedback. There is a precedent for using a data-driven approach to feedback in healthcare more broadly [2, 15, 16]. Studies of data-driven feedback for healthcare providers showed improved team performance and this

approach has been evaluated in both the social science and sporting literature [15–19].

As a result, in this study, we set out to (A) *examine the utility of basic quantitative debriefing performance data collected in real-time*; (B) *to compare the use of this data to existing assessment tools (i.e. DASH)*; and (C) *to assess the future role of this approach for debriefer faculty development* [7, 20].

## Methods

### Study setting

The study was a collaboration between experienced debriefers at the Center for Advanced Pediatric and Perinatal Education (CAPE) at Stanford University (USA) and two Australian SBME centres in the Western Sydney Local Health District network [14]. This study explored the use of recording length of contributions during debriefings and use of conversational diagrams as a means of assessment of debriefing performance with reporting based on STROBE statement guidelines [21].

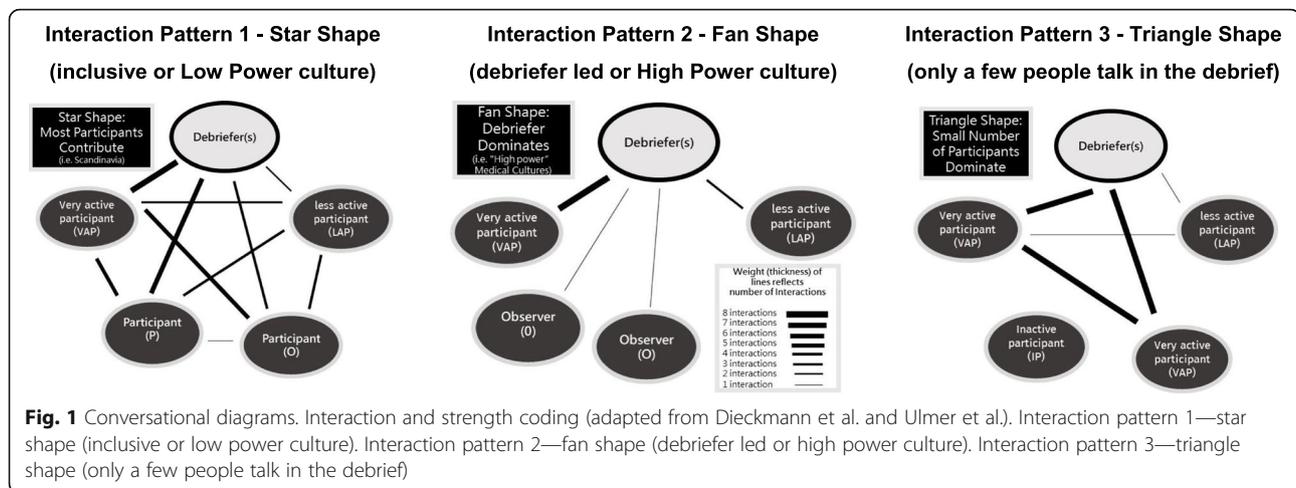
### Inclusion criteria and study subjects

Following the written consent of all simulation participants, debriefers and supervisors, we observed a series of 12 debriefings across two simulation sites. Debriefings were enrolled from January to March 2019 as a convenience sample selected on occasions where the availability of experienced supervising faculty as per the definition by Cheng et al. [13] allowed completion of the study protocols. At the time of data collection, COVID-19 pandemic social distancing restrictions were not in place. Observations and recording were conducted in real time for various elements using a paper data collection sheet. All the debriefings had a single lead debriefer and two supervising faculty present.

### Outcome measures

We recorded the following data points in real time: (A) *study subject interactions* [7] (Fig. 1); (B) *timings*; (C) *quality* (DASH scores) [8] and (D) *demographics*. Demographics included role, gender and debriefing experience. Study subject age was not recorded. Junior doctors were defined as postgraduate year (PGY) 3 or less.

An a priori plan was made to assess the relationship between each member attending the debriefing by hand-drawing conversational diagrams for each debriefing (Fig. 1) [7]. The figures provided reflect the distribution of interactions, timing of each person and the relative strength of the interactions between each study subject. Two investigators observed each of the debriefings. Investigator A recorded the demographics of study subjects while Investigator B measured total time and the duration of conversation that each debriefing study subject contributed. Based on Dieckmann's approach, we



draw a line between two study subjects on the diagram who shared a strong interaction, which is defined as either a question and response or two connected statements in a debriefing [7]. In each resulting diagram, circles show each study subjects their roles and contribution timing, while the lines represent the significant interactions. As the study was in real time, we simplified diagram coding by not separating statements/questions exchanged between each person.

Utterances and gestures were not included in our scoring. Electronic diagrams presented were directly transcribed from free-hand drawn original diagrams. Any freehand or illegible annotations ( $n = 5$ ) noted were excluded from the resultant electronic diagrams.

The timings of contributions of individual study subjects were measured using PineTree Watches™ Version 2.7.0 a multiple subjects stopwatch ([www.pinetreesw.com](http://www.pinetreesw.com)). At the conclusion of each debriefing, Investigator A collected individual DASH scores from study subjects and completed the supervisor version of the scores [8].

#### Debriefers feedback

Following each debriefing, semi-structured feedback was provided from supervising faculty. This was intentionally supplemented by referencing the data collection and was limited to 10 min. The approach used hybridised the feedback methods described by Cheng et al. with the use of timing data and relational diagrams described above [14]. We assessed the impact by asking debriefers for a rating of the usefulness of the information provided (Likert scale 1–5).

#### Analysis plan

Data were analysed using IBM SPSS (V24). Mean and standard deviation (SD) were used to summarise continuous variables. Frequencies and percentages were used for categorical variables. A two-sample t-test was

used to test for differences in the distribution of continuous variables. A gestalt assessment of shape type (Fig. 1) was based on the work of Dieckmann et al. [7].

#### Results

Seventy eight simulation participants were enrolled comprising a mix of students ( $n = 14$ ); doctors ( $n = 54$ ); registered nurses ( $n = 9$ ) and ambulance officers ( $n = 1$ ). There was a high proportion (48.7%) of junior doctors and predominance of female subjects (53.8%). Baseline expertise of debriefers is outlined in Table 1 (divided into novice, intermediate, experienced) based on work by Cheng et al. [13]. The supervising faculty ( $n = 5$ ) were all experienced based on Cheng's work. Figure 2 shows detailed contributions of all simulation participants, debriefers and supervising faculty combined with an illustrative representation of their interactions. The diagrams produced were a mixture of shapes (Fig. 1). In cases where debriefers talked for longer than the participants (ratio of  $\geq 1:1$ ), a fan-shaped appearance was typically observed. This shape is seen in cases 2, 5, 6, 7, 11 and 12 all of which had timing of contribution ratios suggesting relative debriefer 'dominance' (Fig. 2). Cases 1, 4 and 9 had a star-shaped appearance and all had a predominance of contributions from simulation participants (ratio of  $< 1:1$ ). DASH Element 1 simulation participants' ratings in the  $< 1:1$  debriefings were higher than in the remaining ( $\mu = 6.79$  vs  $\mu = 6.44$ ;  $p = 0.036$ ). None of the debriefings displayed a triangular shape, though we observed that students contributed less in large debriefings (i.e. cases 7, 8 and 11). Of note, nursing simulation participants appeared to contribute less to discussions than medical colleagues in the larger interdisciplinary debriefings (i.e. cases 10 and 12).

DASH scores were provided by all simulation participants. For all six elements of the DASH scores, the debriefer self-assessments were much lower than the

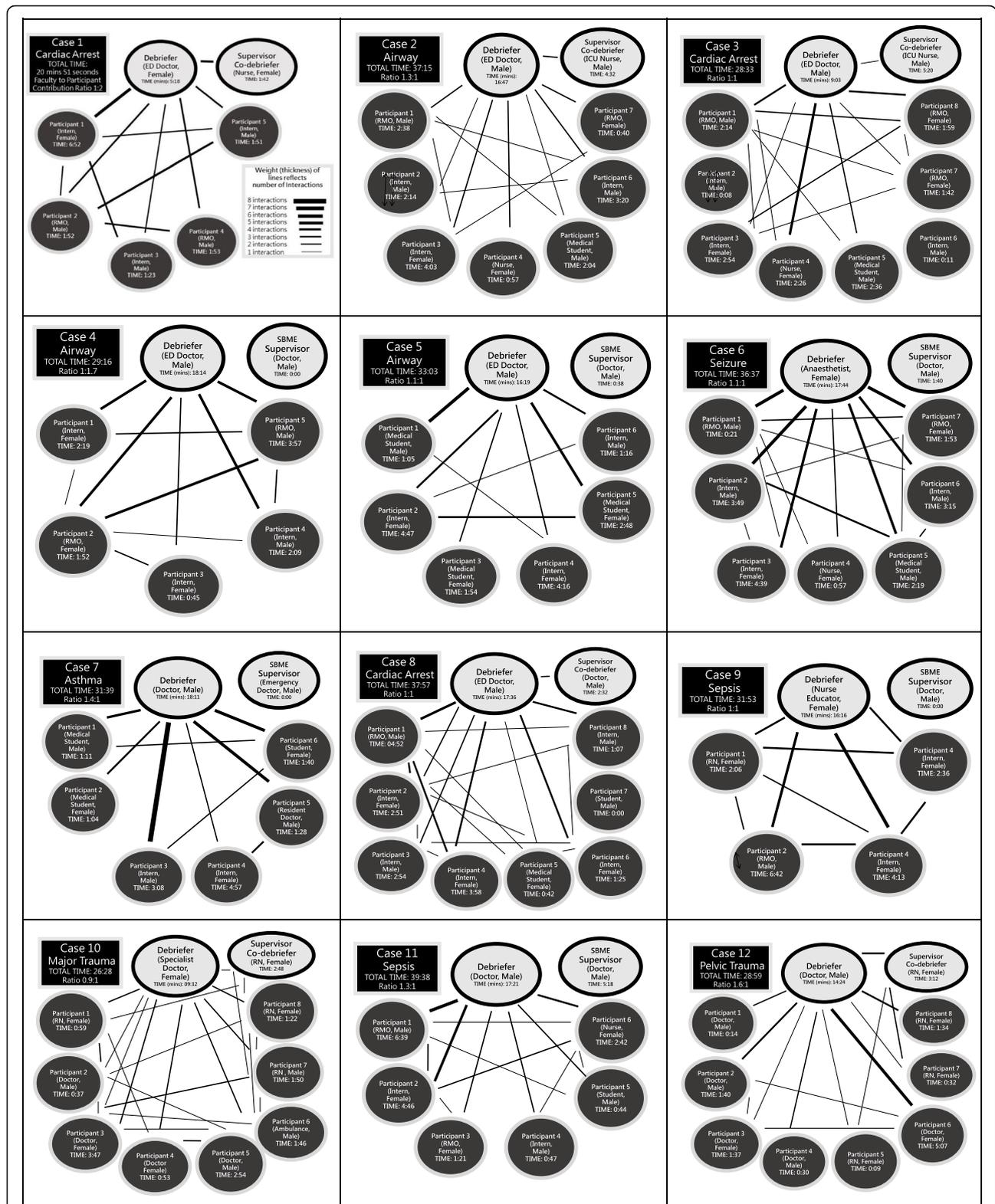
**Table 1** Observational assessment of consecutive SBME debriefings

Topic	Study subject characteristics**	DASH scores					
		DASH element 1 Mean (SD)	DASH element 2 Mean (SD)	DASH element 3 Mean (SD)	DASH element 4 Mean (SD)	DASH element 5 Mean (SD)	DASH element 6 Mean (SD)
Case 1 Adult cardiac arrest	5 study subjects: 5 junior doctors (4 male; 1 female) 1 debriefer: emergency registrar A, (male—novice*)	6.6 (0.49)	6.6 (0.49)	7 (0)	6.8 (0.4)	6 (0.63)	6.4 (0.49)
	1 supervisor: ICU nurse A (female—experienced*)	5	4	5	4	4	5
Case 2 Airway	7 study subjects: (5 junior doctors (3M; 2F), RN (F) student (F)) 1 debriefer: emergency registrar B (female—novice*)	6.71 (0.45)	6.86 (0.35)	6.71 (0.45)	6.57 (0.49)	6.43 (1.05)	6.57 (0.49)
	1 supervisor: ICU nurse B (male—experienced*)	5	6	4	6	6	5
Case 3 Adult cardiac arrest	8 study subjects: (8 junior doctors (4M; 4F)) 1 debriefer: emergency registrar C (male—novice*)	6.63 (0.70)	6.63 (0.70)	6.75 (0.43)	6.25 (0.97)	5.88 (1.45)	6.63 (0.48)
	1 supervisor: ICU nurse B, (male—experienced*)	4	5	6	5	5	6
Case 4 Airway	5 study subjects: (5 junior doctors (3F; 2M)) Debriefers: emergency registrar A (male—intermediate*)	6.8 (0.4)	6.8 (0.4)	7 (0)	7 (0)	7 (0)	7 (0)
	1 supervisor: specialist A (male—experienced*)	5	5	5	4	5	5
Case 5 Airway	6 study subjects: junior doctors (1M, 2F); 3 students (1M, 2F) Debriefers: emergency registrar A (male—intermediate*)	6.17 (0.37)	6.5 (0.5)	6.67 (0.47)	6.67 (0.47)	6 (1.15)	6.5 (0.5)
	1 supervisor: nurse educator A (female—intermediate*)	6	6	7	6	4	5
Case 6 Seizure	7 study subjects: 3 junior doctors (2M; 1F); 4 students (2M; 2F) 1 debriefer: anaesthetic Doctor A (female—intermediate*)	6.57 (0.49)	6.57 (0.49)	6.86 (0.35)	6.86 (0.35)	6.43 (0.73)	6.57 (0.49)
	Supervisor: specialist A (male—experienced*)	5	5	4	5	5	5
Case 7 Asthma	6 Study Subjects: 3 Junior Doctors (2M; 1F), 3 Students (1M; 2F) Debriefers: Emergency Registrar A (Male - Intermediate*)	6.67 (0.47)	7 (0)	7 (0)	7 (0)	6.5 (0.5)	7 (0)
	Supervisor: Specialist (Male - Experienced*)	6	6	7	6	4	5
Case 8 Adult cardiac arrest	8 study subjects: 6 junior doctors (3M; 3F); 2 students (1M; 1F) Debriefers: emergency registrar C (male—intermediate*)	6.75 (0.43)	6.88 (0.33)	7 (0)	6.88 (0.33)	7 (0)	6.88 (0.33)
	Supervisor: specialist (male—experienced*)	6	6	6	5	6	6
Case 9 Sepsis	4 study subjects: 3 junior doctors (1M; 2F), 1 RN (1F) Debriefers: nurse educator A (female—experienced*)	7 (0)	7 (0)	6.75 (0.43)	7 (0)	6.75 (0.43)	7 (0)
	Supervisor: specialist (male—experienced*)	5	5	4	4	4	4
Case 10 Trauma	8 study subjects: 4 doctor (2M; 2F), 3 RN (3F), 1 Ambu (M) Debriefers: emergency specialist A (female—experienced*)	5.75 (0.43)	6.38 (0.48)	5.88 (0.60)	6.13 (0.78)	5 (1.22)	5.63 (0.86)
	Supervisor: nurse (female—experienced*)	4	5	5	4	5	5
Case 11	6 study subjects: 4 doctor (2M; 2F), 1 RN (F), 1 student (M)	6.67 (0.47)	6.67 (0.47)	6.5 (0.5)	6.5 (0.76)	6 (1)	6 (1)

**Table 1** Observational assessment of consecutive SBME debriefings (Continued)

Topic	Study subject characteristics**	DASH scores					
		DASH element 1 Mean (SD)	DASH element 2 Mean (SD)	DASH element 3 Mean (SD)	DASH element 4 Mean (SD)	DASH element 5 Mean (SD)	DASH element 6 Mean (SD)
Sepsis	Debriefer: emergency registrar B (male—intermediate*)	5	4	3	4	4	4
Case 12 Trauma	Supervisor: specialist (male—experienced*)	6	6	5	6	6	5
	8 study subjects: 5 doctor (3M; 2F), 3 RN (3F)	6.13 (0.60)	6.25 (0.66)	6.13 (1.05)	6.38 (0.86)	5.13 (1.45)	5.88 (0.93)
Overall DASH scores	Debriefer: emergency specialist B (male—experienced*)	5	6	5	6	5	5
	Supervisor: nurse (female—experienced*)	6	6	5	6	5	4
	Mean study subject DASH score (n = 78) (mean of total = 6.50)	6.50 (0.34)	6.64 (0.25)	6.66 (0.37)	6.63 (0.30)	6.12 (0.64)	6.46 (0.45)
	Mean self-debriefer DASH score (n = 12)(mean of total = 5.01)	5 (0.58)	5.17 (0.69)	4.83 (0.90)	4.92 (0.86)	5.08 (0.76)	5.08 (0.64)
	p value (t test)	p = < 0.0001					

\*Footnote<sup>13</sup>: < 10 Debriefings = novice ('discovery'), 10–50 Debriefings = intermediate ('growth'), > 50 Debriefings = experienced ('maturity')



**Fig. 2** Timings and interactive diagrams of debriefings (n = 12)

ratings provided by the participants. The differences reached statistical significance for all six DASH elements ( $p < 0.001$ ). In regards to debriefers' experience, of the 12 questionnaires shared 10 were returned resulting in a response rate of 83.3%. Debriefers rating the 'usefulness' of quantitative data provided for their feedback and indicated they found it useful ( $\bar{u} = 4.6/5$  SD 0.49).

## Discussion

Debriefers have the challenging and rewarding task of guiding simulation participants in their post-experience reflection—both by affirming good behaviours and facilitating the remedy of shortfalls in performance [6, 22]. A debriefer's ability to guide participants plays an important role in the delivery of simulation. In this observational study the striking findings included a predominance of debriefers talking more than participants (Fig. 2), significantly higher DASH scores provided by participants compared with those self-rated by debriefers and higher participant DASH scores for the debriefers who talked less. In addition, we observed a high level of debriefer satisfaction in using basic quantitative data (timing and diagrams) as an aid to providing feedback. We have structured the following discussion based on the three objectives outlined in the background section.

### Can real-time quantitative debriefing performance data be used for feedback?

This study assessed the use of timing data and conversational diagrams. Debriefers receiving feedback based on this data rated its 'usefulness' as  $\bar{u} = 4.6$  on a 5-point Likert scale. This is an encouraging finding. While it does not guarantee translation into better debriefing, in other settings data-driven feedback has been shown to significantly improve performance [2, 23]. This study was interrupted by the recent COVID-19 pandemic leading to an under-recruitment of debriefings ( $n = 12$ ), yet we were still able to observe a broad range of interdisciplinary simulation participants and 7 debriefers across 2 SBME sites (Table 1). This suggests that results can be extrapolated to other locations.

Regarding the use of timing data, we present the results for individual times and ratios of contributions of debriefers versus simulation participants (Fig. 2). While the timing data set is interesting within the boundaries of study conditions, it is unclear if this would be practical to measure in typical simulation environments due to resource constraints. It is also unclear whether individual timing data is useful to the debriefers receiving feedback or whether timings reflect quality. For example, knowing an individual talked for a certain period does not necessarily reflect the usefulness of the content, nor the appropriateness of the contribution for promoting

reflection. Within these limitations, in using the data for feedback we found it easy to start meaningful conversations with the debriefers about improving future performance [14]. For example, the data on timing allowed discussion of how to involve quieter participants, and how to increase the number of questions that encouraged reflection rather than 'guess what I am thinking'. While the availability of timings and diagrams appeared to help with feedback, this information arguably may also have been provided using direct observation alone.

From a practical standpoint, we recommend for measuring timing data that a chess clock would be sufficient. A chess clock can provide a simplified binary division of simulation participant and debriefer contributions and would be more practical than our tested method. This approach could provide an estimation of how much the debriefer is talking compared to the participants [6]. With this in mind, from the study findings we note that many debriefings appear to fit a 'sage on the stage' category. This is evidenced by 9/12 debriefings in which facilitators talked for equal or longer than the simulation participants. This important finding may be explained by the increasing requirement of multiple hats to be worn by simulation educators or by a lack of training in our debriefer cohort. Debriefers may revert into more familiar approaches to teaching during debriefings such as tutoring, explanations and lecturing [24]. To address this problem, timing data could help shape future behaviour. Of interest, in a concurrent study we are also investigating the use of cumulative tallies of debriefer questions, debriefer statements and simulation participant responses. In a similar way to using the chess clock approach for timing, this approach may provide an easy to measure method of estimating the debriefer inclusivity.

In regard to the conversational diagrams, these illustrations were used concurrently with the timing data (Fig. 2) for feedback to debriefers. These diagrams were described by Dieckmann et al. in terms of typical roles in SBME, as well as Ulmer et al who described a variety of diagram shapes according to culture [7, 20]. We divided the debriefings in terms of those where the debriefer(s) spoke more than or equal to simulation participants ( $n = 9$ ) and events where the debriefer(s) spoke less ( $n = 3$ ). Using this binary division as a basis for analysis, we observed a pattern in the corresponding shapes of the diagrams (Fig. 2). Similar appearances and shape patterns were reported in Dieckmann and Ulmer's papers [7, 20]. However, on close inspection of each diagram obtained, we could not find the triangular pattern described by Dieckmann et al. The triangle pattern (Fig. 1) is suggestive of 2 or 3 participants dominating. An absence of this pattern was surprising as the range of contribution lengths varied widely (Fig. 2) with some participants not talking at all and some participants talking for  $> 6$  min.

This finding could be due to errors in diagram drawings or random chance. Future studies could avoid any uncertainty in this area by analysing debriefings carefully with use of video and professional transcription.

The astute reader would note that medical students contributed less in larger debriefings (i.e. cases 7, 8 and 11) and nurses contributed less than physicians in mixed groups (i.e. cases 10 and 12). This important observation reminds us of the importance of ensuring a simulation learning climate that feels safe for all, and that the topics chosen for discussion in the debriefing are of interest to all [25–27]. In this study, the majority of recorded interactions were between the debriefer and simulation participants. Very few interactions were recorded between the participants—an important omission—which may represent a target for our own approaches to faculty development at a local level.

In summary, the simulation literature outlines an array of behavioural characteristics exhibited by debriefers that can promote improved future performance [6]. Existing assessment tools such as DASH have an established role in identifying these characteristics. Use of timing data and conversational diagrams may represent an adjunct which may help debriefers understand their performance, track changes over time and assist supervisors in providing debriefer feedback.

#### **How does quantitative debriefing performance data compare to existing tools?**

Existing debriefing assessment tools such as DASH have pros and cons that have been briefly described in the background section. In this study DASH scores were provided by all debriefers and simulation participants. While this was not the primary outcome, it shines a light on the limitations of DASH use. Of note, the 7 debriefers rated themselves significantly lower than the scores from the simulation participants for all DASH elements. These findings reflect our personal experience of using DASH. Prior to the study we had also observed debriefers underscoring themselves and simulation participants overscoring. This finding is interesting, and may be explained by the phenomenon of ‘*response bias*’, which is reported as a problem of assessment scales and surveys [28, 29]. Variation in DASH scores between raters, as well as the time that DASH takes to complete, may reflect the relative subjectivity of the scores provided and limit its value for debriefer feedback [14]. Furthermore, neither the DASH nor OSAD provide specific measurable goals for new debriefers to target in their next debriefing. Therefore, we suggest a continued use of DASH for highlighting ideal behaviours with supplementation of the various quantitative data tools we have outlined in this paper.

#### **What is the potential role of these findings in the development of debriefers?**

As stated, the recipe for success for debriefer faculty development may not come from a single approach. In this study, we found the availability of both quantitative and qualitative data was useful. Experience of using timing data and diagrams together was generally positive, but recording the data and applying this approach was resource intensive. Moreover, the recent pandemic has limited SBME in-person interactions, making current applicability questionable. In the context of the current remote learning climate, a recent study recognised that current methods of faculty development lack a structured approach [30]. We agree that structure is clearly an important factor that faculty development programmes might lack [11]. The quantitative approaches described in our work may assist with providing this structure at the point of care by allocating our attention to observing debriefings in a focused manner. The approaches described should not supercede local planning, adequately resourced and culturally sensitive debriefer faculty development [11, 30].

In terms of other solutions to a relative lack of structure in faculty development programmes, some experts have proposed the use of *DebriefLive*®. This is a virtual teaching environment that allows any debriefer to review their performance [30]. Using this (or similar) software could allow debriefers to observe videos, rate themselves and track progress. In view of the current need for social distancing and the use of remote learning, video review may be an alternative to use of the paper forms and real-time feedback that we used [31–33].

#### **Limitations**

The limitations of our findings are acknowledged especially in relation to the relatively small sample size of the study. We also accept that results are context specific and the approaches described would prove challenging outside of a research setting. Regarding use of the DASH tool as a ‘gold standard’, we note that this tool has been through limited validation. The relevant study used 3 example videos that were scored remotely by online reviewers [8]. On the other hand, validation of OSAD was much broader with studies conducted on electronic versions and in languages other than English [12, 33, 34]. We acknowledge that it is possible our results would have been different had OSAD been used [10]. Regardless, it is our view that the use of any tool as a single approach to faculty development is limited. Locally, we are now using the tools listed above with the data-driven approach assessed in the study [35]. We use either video conferencing or a real-time approach depending on the current local policy on social distancing and remote learning [36].

In conclusion, the use of quantitative data alongside traditional approaches to feedback may be useful for both debriefers looking to improve their future performance and supervising faculty seeking to improve local faculty development programmes.

#### Abbreviations

CAPE: Center for Advanced Pediatric and Perinatal Education; COVID-19: Coronavirus disease 2019; DASH: Debriefing Assessment for Simulation in Healthcare; ED: Emergency department; ICU: Intensive care unit; OSAD: Objective Structured Assessment of Debriefing; PGY: Postgraduate year; RN: Registered nurse; SBME: Simulation-based medical education; STROBE: Strengthening the Reporting of Observational studies in Epidemiology; WSLHD: Western Sydney Local Health District

#### Acknowledgements

The authors would like to thank Nathan Moore, Andrew Baker, Sandra Warburton and Nicole King for supporting the ongoing work of the WSLHD Simulation Centres.

#### Authors' contributions

AC and LM conceived the study. AC and KB extracted data from the simulation centre database and collated the data. AC performed the analysis of results. All authors contributed to and have approved the final manuscript.

#### Funding

The Health Education and Training Institute (HETI) provided limited funding for the simulation equipment used in the simulations through the Medical Education Support Fund (MESF), although this did not directly fund this study.

#### Availability of data and materials

No additional data available.

#### Declarations

##### Ethics approval and consent to participate

The protocols for this study were prospectively examined and approved by the Western Sydney Local Health District (WSLHD) research and ethics committee (HREC 2020/ETH01903).

##### Consent for publication

All authors consent to publication.

##### Competing interests

The authors declare that they have no competing interests.

##### Author details

<sup>1</sup>Department of Emergency Medicine, Westmead Hospital, Sydney, NSW 2145, Australia. <sup>2</sup>Sydney Medical School, Westmead Hospital Block K, Level 6, Westmead Hospital, Sydney, NSW, Australia. <sup>3</sup>Division of Neonatal and Developmental Medicine, Department of Pediatrics, Stanford University, Palo Alto, CA, USA.

Received: 23 September 2021 Accepted: 7 February 2022

Published online: 07 March 2022

#### References

- 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):S52–7. <https://doi.org/10.1097/SIH.0b013e31822724d0>.
- Wolfe H, Zebuhr C, Topjian AA, Nishisaki A, Niles DE, Meaney PA, et al. Interdisciplinary ICU cardiac arrest debriefing improves survival outcomes\*. *Crit Care Med.* 2014;42(7):1688–95. <https://doi.org/10.1097/CCM.0000000000000327>.
- Lee J, Lee H, Kim S, Choi M, Ko IS, Bae J, et al. Debriefing methods and learning outcomes in simulation nursing education: a systematic review and meta-analysis. *Nurse Educ Today.* 2020;87:104345. <https://doi.org/10.1016/j.nedt.2020.104345>.
- Ahmed M, Sevdalis N, Vincent C, Arora S. Actual vs perceived performance debriefing in surgery: practice far from perfect. *Am J Surg.* 2013;205(4):434–40. <https://doi.org/10.1016/j.amjsurg.2013.01.007>.
- Sawyer T, Eppich W, Brett-Fleegler M, Grant V, Cheng A. More Than One Way to Debrief: A Critical Review of Healthcare Simulation Debriefing Methods. *Simul Healthc.* 2016;11(3):209–17. <https://doi.org/10.1097/SIH.0000000000000148>.
- Eppich W, Cheng A. Promoting Excellence and Reflective Learning in Simulation (PEARLS): development and rationale for a blended approach to health care simulation debriefing. *Simul Healthc.* 2015;10(2):106–15. <https://doi.org/10.1097/SIH.0000000000000072>.
- Dieckmann P, Molin Friis S, Lippert A, Ostergaard D. The art and science of debriefing in simulation: Ideal and practice. *Med Teach.* 2009;31(7):e287–94. <https://doi.org/10.1080/01421590902866218>.
- Brett-Fleegler M, Rudolph J, Eppich W, Monuteaux M, Fleegler E, Cheng A, et al. Debriefing assessment for simulation in healthcare (DASH): development and psychometric properties. *Simul Healthc.* 2012;7(5):288–94. <https://doi.org/10.1097/SIH.0b013e3182620228>.
- Rivière E, Aubin E, Tremblay SL, Lortie G, Chiniara G. A new tool for assessing short debriefings after immersive simulation: validity of the SHORT scale. *BMC Med Educ.* 2019;19(1):82. <https://doi.org/10.1186/s12909-019-1503-4>.
- Arora S, Ahmed M, Paige J, Nestel D, Runnacles J, Hull L, et al. Objective structured assessment of debriefing: bringing science to the art of debriefing in surgery. *Ann Surg.* 2012;256(6):982–8. <https://doi.org/10.1097/SLA.0b013e3182610c91>.
- Cheng A, Grant V, Dieckmann P, Arora S, Robinson T, Eppich W. Faculty development for simulation programs: five issues for the future of Debriefing Training. *Simul Healthc.* 2015;10(4):217–22. <https://doi.org/10.1097/SIH.0000000000000090>.
- Runnacles J, Thomas L, Korndorffer J, Arora S, Sevdalis N. Validation evidence of the paediatric objective structured assessment of debriefing (OSAD) tool. *BMJ Simul Technol Enhanc Learn.* 2016;2(3):61–7. <https://doi.org/10.1136/bmjstel-2015-000017>.
- Cheng A, Eppich W, Kolbe M, Meguerdichian M, Bajaj K, Grant V. A conceptual framework for the development of debriefing skills: a journey of discovery, growth, and maturity. *Simul Healthc.* 2020;15(1):55–60. <https://doi.org/10.1097/SIH.0000000000000398>.
- Cheng A, Grant V, Huffman J, Burgess G, Szyld D, Robinson T, et al. Coaching the debriefer: peer coaching to improve debriefing quality in simulation programs. *Simul Healthc.* 2017;12(5):319–25. <https://doi.org/10.1097/SIH.0000000000000232>.
- Halamek L, Cheng A. Debrief to Learn Edition 9 - NASA debriefing methods. <https://debrief2learn.org/podcast-009-nasa-debriefing-methods> (accessed 4 June 2021)
- 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(10):2817–22. <https://doi.org/10.1097/CCM.0b013e318186fe37>.
- Santos-Fernandez E, Wu P, Mengersen KL. Bayesian statistics meets sports: a comprehensive review. *Journal of Quantitative Analysis in Sports.* 2019;15(4):289–312. <https://doi.org/10.1515/jqas-2018-0106>.
- Evenson A, Harker PT, Frei FX. Effective Call Center Management: Evidence from Financial Services. Philadelphia: Wharton School Center for Financial Institutions, University of Pennsylvania; 1999.
- Fernández-Echeverría C, Mesquita I, González-Silva J, Moreno MP. Towards a More Efficient Training Process in High-Level Female Volleyball From a Match Analysis Intervention Program Based on the Constraint-Led Approach: The Voice of the Players. *Front Psychol.* 2021;12(563):645536.
- Ulmer FF, Sharara-Chami R, Lakissian Z, Stocker M, Scott E, Dieckmann P. Cultural prototypes and differences in simulation debriefing. *Simul Healthc.* 2018;13(4):239–46. <https://doi.org/10.1097/SIH.0000000000000320>.
- Cheng A, Kessler D, Mackinnon R, Chang TP, Nadkarni VM, Hunt EA, et al. Auerbach M; International Network for Simulation-based Pediatric Innovation, Research, and Education (INSPIRE) Reporting Guidelines Investigators. Reporting Guidelines for Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements. *Simul Healthc.* 2016;11(4):238–48. <https://doi.org/10.1097/SIH.0000000000000150>.
- Tekian A, Watling CJ, Roberts TE, Steinert Y, Norcini J. Qualitative and quantitative feedback in the context of competency-based education. *Med*

- Teacher. 2017;39(12):1245–9. <https://doi.org/10.1080/0142159X.2017.1372564>.
23. Johnson M, Peat A, Boyd L, Warren T, Eastwood K, Smith G. The impact of quantitative feedback on the performance of chest compression by basic life support trained clinical staff. *Nurse Educ Today*. 2016;45:163–6. <https://doi.org/10.1016/j.nedt.2016.08.006>.
  24. Harden RM, Lilley P. The Purpose and Function of a Teacher in the Healthcare Professions. In: *The Eight Roles of the Medical Teacher*. Amsterdam: Elsevier Health Sciences; 2018.
  25. Rudolph JW, Raemer DB, Simon R. Establishing a safe container for learning in simulation: the role of the pre-simulation briefing. *Simul Healthc*. 2014; 9(6):339–49. <https://doi.org/10.1097/SIH.0000000000000047>.
  26. Salik I, Paige JT. Debriefing the Interprofessional Team in Medical Simulation. [Updated 2021 Apr 28]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. <https://www.ncbi.nlm.nih.gov/books/NBK554526>. Accessed 25 July 2021
  27. Tavares W, Eppich W, Cheng A, Miller S, Teunissen PW, Watling CJ, et al. Learning conversations: an analysis of their theoretical roots and their manifestations of feedback and debriefing in medical education. *Acad Med*. 2020; <https://doi.org/10.1097/ACM.0000000000002932>.
  28. Morey JC, Simon R, Jay GD, Wears RL, Salisbury M, Dukes KA, et al. Error reduction and performance improvement in the emergency department through formal teamwork training: evaluation results of the MedTeams project. *Health Serv Res*. 2002;37(6):1553–81. <https://doi.org/10.1111/1475-6773.01104>.
  29. Kreitchmann RS, Abad FJ, Ponsoda V, Nieto MD, Morillo D. Controlling for response biases in self-report scales: forced-choice vs. psychometric modeling of Likert items. *Front Psychol*. 2019;10:2309. Published 2019 Oct 15.
  30. Wong NL, Peng C, Park CW, Pérez J 4th, Vashi A, Robinson J, et al. DebriefLive: A pilot study of a virtual faculty development tool for debriefing. *Simul Healthc*. 2020;15(5):363–9. <https://doi.org/10.1097/SIH.0000000000000436>.
  31. Cheng A, Kolbe M, Grant V, Eller S, Hales R, Symon B, et al. A practical guide to virtual debriefings: communities of inquiry perspective. *Adv Simul (Lond)*. 2020;5:18.
  32. Abegglen S, Krieg A, Eigenmann H, Greif R. Objective Structured Assessment of Debriefing (OSAD) in simulation-based medical education: translation and validation of the German version. *PLoS One*. 2020;15(12):e0244816. <https://doi.org/10.1371/journal.pone.0244816>.
  33. Zamjahn JB, Baroni de Carvalho R, Bronson MH, Garbee DD, Paige JT. eAssessment: development of an electronic version of the Objective Structured Assessment of Debriefing tool to streamline evaluation of video recorded debriefings. *J Am Med Inform Assoc*. 2018;25(10):1284–91. <https://doi.org/10.1093/jamia/ocy113>.
  34. Chung HS, Dieckmann P, Isenberg SB. It is time to consider cultural differences in debriefing. *Simul Healthc*. 2013;8(3):166–70. <https://doi.org/10.1097/SIH.0b013e318291d9ef>.
  35. Husebø SE, Dieckmann P, Rystedt H, Søreide E, Friberg F. The relationship between facilitators' questions and the level of reflection in postsimulation debriefing. *Simul Healthc*. 2013;8(3):135–42. <https://doi.org/10.1097/SIH.0b013e31827cbb5c>.
  36. Monette DL, Macias-Konstantopoulos WL, Brown DFM, Raja AS, Takayesu JK. A video-based debriefing program to support emergency medicine clinician well-being during the COVID-19 pandemic. *West J Emerg Med*. 2020;21(6):88–92. <https://doi.org/10.5811/westjem.2020.8.48579>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

