Development of the Human Factors Skills for Healthcare Instrument: a valid and reliable tool for assessing interprofessional learning across healthcare practice settings

Gabriel B Reedy,[1,2] Mary Lavelle,[1,3] Thomas Simpson,[1,2] Janet E Anderson[3]

OPEN ACCESS

ABSTRACT

Background A central feature of clinical simulation training is human factors skills, providing staff with the social and cognitive skills to cope with demanding clinical situations. Although these skills are critical to safe patient care, assessing their learning is challenging. This study aimed to develop, pilot and evaluate a valid and reliable structured instrument to assess human factors skills, which can be used pre- and post-simulation training, and is relevant across a range of healthcare professions.

Method Through consultation with a multi-professional expert group, we developed and piloted a 39-item survey with 272 healthcare professionals attending training courses across two large simulation centres in London, one specialising in acute care and one in mental health, both serving healthcare professionals working across acute and community settings. Following psychometric evaluation, the final 12-item instrument was evaluated with a second sample of 711 trainees.

Results Exploratory factor analysis revealed a 12-item, one-factor solution with good internal consistency (α=0.92). The instrument had discriminant validity, with newly qualified trainees scoring significantly lower than experienced trainees (t(98)=4.88, p<0.001) and was sensitive to change following training in acute and mental health settings, across professional groups (p<0.001). Confirmatory factor analysis revealed an adequate model fit (RMSEA=0.066).

Conclusion The Human Factors Skills for Healthcare Instrument provides a reliable and valid method of assessing trainees' human factors skills self-efficacy across acute and mental health settings. This instrument has the potential to improve the assessment and evaluation of human factors skills learning in both uniprofessional and interprofessional clinical simulation training.

INTRODUCTION

The provision of effective healthcare is a complex choreography involving many professional groups, with diverse cultures, background and language. Individuals must work collaboratively together, delivering a high level of skill in what can be stressful situations. The technical and clinical expertise required is underpinned by the social and cognitive skills needed to cope with the demands of the situation, including situational awareness, communication, teamwork, leadership, decision making and

care—for the self, colleagues and patients.[1 2] These skills are often referred to as non-technical skills; however, this term is problematic. The distinction between technical and non-technical skills is not the same across disciplines in healthcare. For example, in mental health, communication is the medium through which assessments are made, diagnoses are formed and treatment is delivered, thus making communication a technical skill in this context.[3] Nestel *et al* argue that the distinction between technical and non-technical skills is unhelpful, and that successful patient care in the dynamic complex system of healthcare requires integration of a broad range of skills.[3] For these reasons, we use the term 'human factors skills for healthcare', drawing on the disciplinary knowledge of human factors to examine the social and cognitive skills relevant across a range of healthcare professions.

Human factors skills deficiencies are a leading cause of clinical error,[4 5] and the value of training these skills is well recognised, specifically improving patient safety[6] and reducing mortality.[7] However, methods of evaluating the effectiveness of training for practitioners in this field are not well developed. A recent meta-analysis of simulation training revealed that the majority of trainee assessments focused on discrete technical skills, such as time to task completion, process measures and task outcomes.[8] Despite some advances in observer-rated assessment in specific specialities such as anaesthesia[9] or surgery,[10] there remains no valid, reliable tool for assessing the knowledge and learning of human factors skills in healthcare training, which limits development and evaluation of training interventions. Even in areas as well defined and relatively experienced in simulation as anaesthesia, the absence of a gold-standard assessment tool for non-technical skills learning has been identified as a major hindrance to the effective evaluation of this modality.[11]

The overarching aim of our simulation centre is to deliver interprofessional training to healthcare staff that improves trainees' understanding of, and competence in, the human factors skills relevant to healthcare. Our approach employs a learner-led debriefing model, the Debrief Diamond,[12] to explore trainees' assumptions and understanding of their own actions in simulated and actual clinical practice. To this end, the learning that we hope trainees will acquire cannot be adequately measured

135

through performance on discrete technical tasks during a specific scenario, but requires a more holistic evaluative approach. Our philosophy of learner-led debriefing has, in turn, led to us developing an evaluation strategy based on learner-reported self-efficacy in human factors skills.

Self-efficacy has been defined as an individual's belief in their own capacity to execute behaviours necessary to perform specific tasks.[13] Self-efficacy measures offer an opportunity to assess individuals' beliefs in their own capacity to execute behaviours necessary to perform specific tasks.[14] Factors that influence self-efficacy are also core features of simulation training and include task performance, observation and modelling of others, verbal encouragement and feedback; physiological components are also an influence, such as stress or anxiety.[15 16] Previous studies have shown that self-efficacy measures are sensitive to change following psychosocial interventions in healthcare contexts[17] including simulation training,[18–20] and correlate with actual performance.[21 22] As such, we believe that simulation-based training provides an ideal learning environment for improving self-efficacy in healthcare human factors skills, and further, that self-efficacy is a good proxy measure for the development of individuals' in situ skills.

The aim of this study was to develop, pilot and evaluate a structured instrument to assess self-efficacy of human factors skills that can be used beforesimulation and after simulation training, and is relevant to a range of healthcare professions.

METHODS
Setting
The study took place in two large simulation centres in South London: The Simulation and Interactive Learning (SaIL) Centre at Guys and St. Thomas' (GSTT) NHS Foundation Trust, which provides simulation training for healthcare professionals working primarily in acute clinical settings and Maudsley Simulation at South London and Maudsley NHS Foundation Trust, providing simulation training for mental health practitioners working in community and acute settings. These clinical simulation facilities are, respectively, located on the campuses of two large inner-London hospitals. Ethical approval for this work was provided by King's College London ethics committee (RESCMR-15/16–1561).

Participants
The instrument was developed and evaluated in two separate phases. During the instrument development phase, the 39-item instrument was completed by 272 healthcare staff attending 17 simulation training courses at the two sites within a 3-month period (April through June 2016). Once the final instrument had been developed, it was then evaluated. During the evaluation phase, the finalised 12-item instrument was completed by 711 healthcare professionals attending simulation training within a 6-month period (September through February 2017). Individuals who participated in the development phase did not participate in the evaluation phase. The participant sample across both phases reflected the interprofessional nature of the training, consisting of healthcare assistants, nurses, doctors and allied healthcare professionals across a range of experience grades and clinical areas.

Procedure
The instrument was completed by participants as part of routine simulation training. Participants were attending 17 different training courses, delivered across two simulation centres.

Although a key component of all training was to improve participants' human factors skills within the context of healthcare, the specific courses focused on a variety of clinical topics (e.g., sepsis, emergency care, psychosis, care for the elderly)) across a range of healthcare settings (eg, acute wards, psychiatric wards, paramedic transportation and care at home). The training was delivered by experienced trainers and clinical educators in each centre. On arrival to the training, participants were informed about the research study both verbally and through a participant information sheet. The instrument was completed by consenting participants twice: once at the start of the training day (pre-course) and for a second time at the end of the training day (post-course). Data collection procedures were identical during the instrument development (April through June 2016) and evaluation (September 2016 through February 2017) phases.

Measures
Participants' sociodemographic information was assessed through a self-report questionnaire, recording gender, age, professional background and the number of years professionally qualified.

Human Factors Skills for Healthcare instrument development
Item generation
The core human factors skills that underpin clinical working were identified from the literature as teamwork, leadership, communication, situational awareness, decision-making and care (including self-care, care and compassion for patients and care and compassion for colleagues).[1 2] Using the six core human factors categories as a framework for the instrument, a multi-professional research team (i.e., acute care clinicians, psychiatrists, educational psychologists and human factors experts) generated a pool of 240 items reflecting these categories. After the initial pool had been generated, the research team reviewed the items for relevance and redundancy, resulting in a reduced pool of 120 items (20 items per category).

To assess the relevance of the item content beyond the research team, the items were presented, via an online survey, to expert colleagues for validation. These individuals were recruited via personal contacts of the research team and through dissemination at international simulation education conferences. The survey was completed by 30 participants from the UK, the USA, Australia and Europe, including doctors, nurses, allied healthcare professionals, experts in fields of human factors, experts in clinical simulation and clinical educators. All responses were equally weighted as all respondents were trained in human factors.

Participants were asked to rate each item for suitability for inclusion on a final instrument, on a four-point scale: 1, definitely include item; 2, possibly include; 3, possibly remove item and 4, definitely remove item. The four-point scale was chosen over the dichotomous categorisation of include/exclude, allowing for greater sensitivity in responses. The survey design also provided an opportunity to suggest potential edits for each item.

Of the 120 items, 10 were unanimously coded as 'definitely include'. These 10 items were all included in the instrument. All the other included items were rated favourably with 75% of the respondents rating the items 'definitely' or 'possibly include'. No item was unanimously coded as 'definitely remove'. Items with definitely include and possibly include rates greater than the mean inclusion rate of the human factors category were selected for inclusion in the instrument (n=29). As such, the

pilot instrument consisted of 39 items across the six human factors categories.

The 39-item pilot instrument included the stem question 'Please rate how confident you are that you can manage the following effectively', consistent with theory and practice in self-efficacy instrument design,[14] and participants were asked to respond on a scale from one to ten.

Statistical analysis
Item selection
Initial analyses were conducted using IBM SPSS (V.22) software.[23] The item selection process included five steps:

1. Participant responses were descriptively explored, investigating ceiling and floor effects.
2. As sensitivity to change pre-training and post-training was an important feature of the instrument, paired samples t-tests assessed the change in item scores pre-course to post-course, Cohen's d effect sizes were calculated for each item and items with a small effect size ($d<0.5$) were eliminated from the instrument.
3. The remaining items were scrutinised by the research teams' educational psychologist (GR), human factors expert (JEA) and clinician (TS) in terms of their relevance and wording. Items that were deemed less relevant or inappropriately worded were removed.
4. Relationships between remaining items were explored using a bivariate Pearson's correlation coefficient to assess for multi-collinearity between items.
5. An exploratory factor analysis (EFA) using a maximum likelihood factor extraction method was conducted. Only factors with Eigenvalues over 1 were extracted.

Sensitivity to change and discriminant validity
The mean total score of the final instrument was calculated for the pre-course and post-course data. Paired samples t-tests compared mean scores of the final instrument pre-training and post-training. Participants were compared overall, by centre and by professional group. To assess the discriminant validity of the instrument, independent samples t-tests compared pre-course scores for newly qualified trainees (qualified 1 year or less) with experienced trainees (qualified 10 years or more).

Instrument evaluation
On the basis of the EFA results, the final 12-item instrument was piloted with 711 healthcare professionals attending simulation training during a six-month period (September 2016 through February 2017). Confirmatory factor analysis (CFA) was conducted on the pre-course response data collected during this time to test the factor structure identified in the EFA. The CFA was conducted in the Analysis of Moment Structures (AMOS) V.22 software.[23]

RESULTS
Sample
Sociodemographic details of participants recruited in the development phase (n=272) and those recruited in the evaluation phase (n=711) are displayed in table 1. Participants in the development phase were qualified for an average of 8.74 years (SD=8.72; range=0–35; median=6 years), and participants recruited in the evaluation phase were qualified for an average of 7.14 years (SD=8.43; range 0–50; median=4 years).

Item selection
The feedback from training participants and simulation trainers was that the 39-item instrument was too long and was not

Table 1 Sociodemographic information of participants in the development and evaluation phases

	Development phase (n=272), n (%)	Evaluation phase (n=711), n (%)
Gender		
Female	207 (76)	580 (82)
Ethnicity		
White	171 (63)	423 (59)
Black/Black British	25 (9)	127 (18)
Asian/Asian British	16 (6)	91 (13)
Mixed	9 (3)	20 (3)
Other ethnic background	12 (4)	50 (7)
Age, years		
<25	30 (11)	158 (22)
25–34	128 (47)	311 (44)
35–45	68 (25)	144 (20)
>45	46 (17)	98 (14)
Profession		
Nurses/Midwives	152 (56)	400 (56)
Doctors	64 (24)	220 (31)
Allied health professionals	56 (21)	91 (13)

practical for use beyond the pilot phase. Trainers suggested that the instrument would need to be substantially reduced to make it a feasible tool to accompany simulation training courses in the simulation centres. With this in mind, we began the five-step item selection process. The process consisted of the following steps:

Step 1: All 39 items showed normal distributions as the skewness of each item remained within normal levels (range: −1.0 to −0.12); thus, no ceiling or floor effects were observed.

Step 2: Paired samples t-test assessed the change in item scores pre-course to post-course, and Cohen's d effect sizes were calculated for each item. Sixteen items with small effect sizes ($d<0.5$) were identified and removed (for details see online supplementary appendix 1).

It is possible that the small effect sizes could be due to limitations of the training rather than lack of sensitivity of the items. However, the data were collected across 17 different training courses at two sites, delivered by a range of trainers and clinical educators (n=20+) on a variety of topics. The diversity of this training, combined with the relatively large sample size, suggests that the small effect size of these items is more likely to be a feature of the items rather than the training.

Step 3: Alongside the quantitative data, piloting the instrument provided informal feedback on how the items were received and interpreted by participants. Incorporating this feedback alongside reflections of the research team, 10 items were removed from the instrument for a number of reasons: five items used ambiguous language (i.e., 'recognising when there is a need for workload re-distribution among colleagues', 'outwardly labelling specific clinical situations', 'awareness of availability of resources in your environment', 'implementing immediate coping strategies to manage your own stress in a busy clinical environment', 'expressing uncertainty to patients and relatives'); two items required participants' metacognition (e.g., 'Recognising how your leadership style impacts others', 'recognising how your state of mind can influence those around you'); one item was less relevant across a range of professional groups ('Summarising patient information to hand over using a tool such as SBAR'), and two items were deemed too vague ('awareness of your own

Table 2 Pearson's correlations (*r*) between items during the development phase are displayed alongside total correlations, alpha if deleted and factor loadings for each item for both the development phase (*n*=272) and the evaluation phase (*n*=711) samples

Items	1	2	3	4	5	6	7	8	9	10	11	12	13	Development phase (n=272) Corrected item-total correlations	alpha if item deleted	Factor Loadings EFA	Evaluation phase (n=711) Corrected item-total correlations	alpha if item deleted	Factor loadings CFA
1. Constructively managing others' negative emotions at work	–	0.51	0.45	0.32	0.33	0.35	0.39	0.36	0.44	0.47	0.42	0.37	0.47	0.55	0.92	0.55	0.69	0.96	0.49
2. Communicating effectively with a colleague with whom you disagree		–	0.39	0.38	0.39	0.38	0.42	0.42	0.40	0.46	0.45	0.32	0.46	0.56	0.92	0.58	0.69	0.96	0.55
3. Prioritising when many things are happening at once			–	0.36	0.40	0.33	0.48	0.36	0.36	0.44	0.44	0.44	0.43	0.55	0.92	0.64	0.80	0.96	0.67
4. Requesting help from colleagues in other professions				–	0.46	0.50	0.50	0.43	0.40	0.41	0.38	0.46	0.52	0.57	0.92	0.55	0.83	0.95	0.74
5. Speaking up as part of a team to convey what you think is going on					–	0.66	0.66	0.48	0.45	0.55	0.54	0.58	0.57	0.69	0.92	0.74	0.84	0.95	0.78
6. Involving colleagues in your decision making process						–	0.64	0.36	0.36	0.55	0.54	0.58	0.54	0.67	0.92	0.71	0.78	0.96	0.70
7. Dealing with uncertainty in your decision making process							–	0.54	0.50	0.56	0.58	0.55	0.63	0.73	0.92	0.75	0.82	0.95	0.71
8. Asking other team members for the information I need during a busy clinical environment								–	0.61	0.53	0.58	0.53	0.55	0.67	0.92	0.73	0.80	0.96	0.69
9. Recognising when you should take on a leadership role									–	0.64	0.64	0.53	0.63	0.68	0.92	0.73	0.83	0.95	0.69
10. Monitoring the 'big picture' during a complex clinical situation										–	0.72	0.65	0.72	0.77	0.92	0.81	0.84	0.95	0.66
11. Anticipating what will happen next in clinical situations											–	0.69	0.63	0.76	0.92	0.82	0.79	0.96	0.57
12. Working effectively with a new team in clinical situations												–	0.57	0.72	0.92	0.75	0.79	0.96	0.63
13. Asking others to take on tasks within the team*													–	0.78	0.92	–	–	–	–

*Item not included in final instrument.
CFA, confirmatory factor analysis; EFA, exploratory factor analysis.
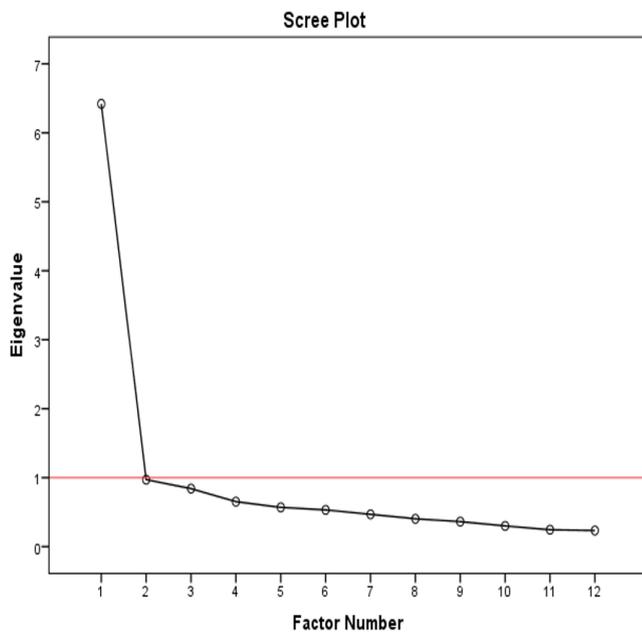
Figure 1 Exploratory factor analysis Scree plot; one-factor model was selected.

behaviour in a clinical setting', 'making critical decisions under pressure').

Step 4: Pearson's correlation coefficients for the remaining 13 items are displayed in table 2, alongside the item-total correlations and the alpha if item deleted. Reliability analysis of the 13 items revealed a Cronbach's alpha of 0.92. All items were significantly positively correlated with inter-item correlations ranging from $r=0.3$ to $r=0.7$. Item 13 (*Asking others to take on tasks within the team*) showed the highest correlations with other items (item total correlation=0.78) with no change in alpha if removed. To reduce redundancy, this item was removed prior to factor analysis.

Step 5: An EFA of pre-course responses was conducted using a maximum likelihood method extracting factors with Eigenvalues greater than 1. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.93, and the Bartlett's test of sphericity was highly significant ($\chi^2=1382.51$, df=66, p<0.0001). According to these two tests, the correlations and partial correlations between the items imply the existence of latent factors and justify the choice to apply an EFA. The Scree plot is displayed in figure 1. This revealed a one-factor solution that explained 53.5% of the variance. The factor loadings of each item are displayed in table 2

(range: 0.55–0.82). The instrument showed good internal consistency (Cronbach's alpha=0.92).

Sensitivity to change
Paired samples *t*-test comparisons of pre-training and post-training scores for the final 12-item instrument are displayed in table 3. The instrument demonstrated sensitivity to change post-training at both acute and mental health training centres and across all professional groups (doctors, nurses and allied health professionals). At every level of analysis, participants showed significant improvement following training ($p<0.001$) with large effect sizes (*d* range: 0.66–0.75) (see table 3).

Discriminant validity
Newly qualified trainees' pre-course scores ($n=52$, $M=6.73$, $SD=1.17$) were significantly lower than experienced trainees' scores ($n=48$, $M=7.84$, $SD=1.07$) $t(98)=4.88$, $p<0.0001$.

Confirmatory factor analysis
During the evaluation phase, the final 12-item instrument was completed by 711 participants attending simulation training during a six-month period (Sep 2016 to Feb 2017). Reliability analysis revealed excellent internal consistency (Cronbach's alpha=0.96) (see table 2).

There are differing opinions on the model design of a CFA. The EFA clearly demonstrated a one-factor model. An unmodified CFA model did not yield an adequate fit ($RMSEA=0.12$). However, Gerbing and Anderson,[24] and more recently Brown,[25] suggest that it is sometimes appropriate to correlate items in the model based on theoretical or methodological grounds, particularly when items are conceptually linked. Doing so may improve the fit, but should be done with caution and should be based on methodological rather than statistical justification.

We argue that human factors concepts are strongly conceptually related to each other[1] and that even with strong correlations it is important to include and to attempt to measure these distinct but related concepts. Therefore, based on these methodological grounds, we fit a one-factor CFA model with correlations between conceptually similar items. Specifically, correlations were added between items one and three, which both refer to the management of difficult situations and between items 9 and 12, which refer to the higher level cognitive aspects of situational awareness and team work. The resulting CFA model revealed an adequate fit for the unidimensional model ($SRMR=0.02$; $RMSEA=0.066$; $TLI=0.97$; $CIF=0.98$).

DISCUSSION
The primary outcome of this study was the development of the Human Factors Skills for Healthcare Instrument (HuFSHI),

Table 3 Paired samples t-test comparisons of mean Human Factors Skills for Healthcare instrument scores by training centre and profession

	Pre-training, M (SD)	Post-training, M (SD)	t	df	p	d
All participants	7.14 (1.2)	7.99 (1.27)	14.60	252	<0.001	0.69
Training centre						
Acute clinical training	7.15 (1.19)	7.98 (1.28)	11.95	194	<0.001	0.67
Mental health training	7.13 (1.26)	8.03 (1.28)	9.26	57	<0.001	0.72
Professional group						
Nursing and midwifery	7.30 (1.21)	8.15 (1.34)	10.42	143	<0.001	0.66
Doctors	7.00 (1.06)	7.75 (1.10)	7.52	59	<0.001	0.68
Allied health professionals	6.86 (1.31)	7.81 (1.22)	7.05	47	<0.001	0.75

which is a 12-item, unidimensional measure of human factors skills self-efficacy applicable to healthcare practitioners. The results show this instrument to be reliable, with face and content validity and the ability to discriminate between newly qualified and experienced healthcare providers. Furthermore, it is sensitive to change post-simulation training and relevant for multiple healthcare professionals (ie, doctors, nurses and allied health professionals) in both acute and mental health training centres that train colleagues in in-patient and community settings.

To the authors' knowledge, this is the first instrument that can be used to evaluate the learning of human factors skills in healthcare across a range of professional groups and across healthcare settings. This is of particular salience given the recent shift in clinical education to support the delivery of more integrated interprofessional training,[26] reflecting the reality of healthcare provision. Feedback from the simulation trainees and trainers participating in the pilot highlighted the need for brevity in the instrument. The competing demands, priorities and time pressures on a simulation training course meant it was difficult to incorporate a 39-item instrument into the training day. As such, we endeavoured to create a reliable and valid brief instrument that could feasibly be used in conjunction with simulation training to assess learning.

The final instrument demonstrated good internal consistency. Content and face validity of the instrument was ensured by the input of the multi-professional team in development and critical appraisal of the items. Attempts were made to make the items task-specific and reflective of the healthcare setting. Discriminant validity was evidenced by the significantly lower pre-training scores of newly qualified trainees, compared with experienced healthcare trainees.

This instrument was developed to assess the six human factors skills: situational awareness, communication, teamwork, leadership, decision-making and care.[1 2] The instrument was unidimensional and factor analysis indicated a high degree of relatedness between all items, which we adjusted for the CFA model. Human factors skills are sometimes termed Non-Technical Skills. This is an umbrella term used to describe related behaviours that influence safe and efficient task execution.[1] Our study provides evidence to confirm the inter-relatedness of these skills.

The skill 'care' was the lowest loading item in the final model, suggesting it has some conceptual distinctiveness. The concept of 'care' in this context is multifaceted and complex, referring to care for the self, for colleagues and for patients. The items reflecting care were challenging to word as they conveyed a sense of social desirability; these items were also less sensitive to change post-training. As such, only one item in the final instrument (item one) refers to care. Thus, perhaps the concept of care in this context could, in the future, be teased apart and explored separately from other human factors skills. This is one avenue for future work in the field.

The findings should be considered in light of the study limitations. First, training sites were both based in South London, and as such the relevance of this instrument to healthcare providers outside inner-city healthcare settings is untested. However, as the instrument framework builds on a theoretical understanding of the core human factors skills identified in literature and is grounded in self-efficacy theory, we would expect this to be applicable to a range of healthcare settings in other areas. Second, as discussed previously, developing unbiased, appropriate items for the human factors category 'care' was challenging, and as a result, this skill may be under-represented in the final instrument. Third, due to a lack of available tools to measure human factors skills

in healthcare, we were unable to assess the criterion validity of the instrument. Through a programme of continued research in this area, we aim to build on our current work addressing some of these limitations. We are currently undertaking an international validation of the instrument, which includes an additional sub-set of 'care' items. Furthermore, to explore the concurrent and predictive validity of the instrument, we aim to use observational tools to assess behavioural markers of human factors skills during simulated scenarios. Ultimately, this line of research could identify proxy measures of actual human factors skills in clinical settings.

CONCLUSION

Simulation-based education often focuses on the human factors skills that underpin much of clinical practice, and yet they remain difficult to assess and evaluate. This makes it difficult to determine whether, as a result of specific simulation training interventions, individual participants have developed in terms of their human factors skills; likewise, it is difficult to evaluate whether particular training interventions are accomplishing their educational aims. Perhaps ideally, social scientists would seek to follow healthcare practitioners back into clinical practice, to trace long-lasting behavioural changes that we hope will have a positive impact on patient safety and care. However, this kind of evidence is still difficult to obtain, largely because the practical implications of collecting such data make it nearly impossible. Until other measures are available, we argue that theoretically grounded, valid and reliable proxy measures can help us make sense of how and whether learners develop these important skills in simulation. Much work remains as we try and understand the nature of learning that occurs in simulation-based settings. We have argued that the HuFSHI provides a reliable and valid method of evaluating trainees' self-efficacy as regards human factors skills—and thus is a particularly useful instrument to help us understand how simulation helps them develop their human factors learning across both acute and mental health settings.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Flin RH, O'Connor P, Crichton M. *Safety at the sharp end: a guide to non-technical skills*. Farnham, UK: Ashgate Publishing, Ltd, 2008.
2 Ensemble C. Performing medicine: circle of Care. secondary performing medicine: circle of Care 2016. www.performingmedicine.com/project/circle-of-care.
3 Nestel D, Walker K, Simon R, *et al*. Nontechnical skills: an inaccurate and unhelpful descriptor? *Simul Healthc* 2011;6:2–3.
4 Kohn LT, Corrigan JM, Donaldson MS. *To err is human: building a safer health system*. Washington, DC: National Academies Press, 2000.
5 Leonard M, Graham S, Bonacum D. The human factor: the critical importance of effective teamwork and communication in providing safe care. *Qual Saf Health Care* 2004;13(Suppl 1):i85–90.
6 McCulloch P, Mishra A, Handa A, *et al*. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. *Qual Saf Health Care* 2009;18:109–15.
7 Neily J, Mills PD, Young-Xu Y, *et al*. Association between implementation of a medical team training program and surgical mortality. *JAMA* 2010;304:1693–700.
8 Cook DA, Hatala R, Brydges R, *et al*. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 2011;306:978–88.
9 Fletcher G, Flin R, McGeorge P, *et al*. Anaesthetists' Non-Technical skills (ANTS): evaluation of a behavioural marker system. *Br J Anaesth* 2003;90:580–8.
10 Tsuburaya A, Soma T, Yoshikawa T, *et al*. Introduction of the non-technical skills for surgeons (NOTSS) system in a japanese Cancer center. *Surg Today* 2016;46:1451–5.
11 Lorello GR, Cook DA, Johnson RL, *et al*. Simulation-based training in anaesthesiology: a systematic review and meta-analysis. *Br J Anaesth* 2014;112:231–45.
12 Jaye P, Thomas L, Reedy G. 'The Diamond': a structure for simulation debrief. *Clin Teach* 2015;12:171–5.
13 Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977;84:191–215.
14 Bandura A. Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents*. , 2006:5, 307–37.
15 Bandura A. Self-efficacy mechanism in human agency. *Am Psychol* 1982;37:122–47.
16 Strecher VJ, DeVellis BM, Becker MH, *et al*. The role of self-efficacy in achieving health behavior change. *Health Educ Q* 1986;13:73–92.
17 Lorig KR, Ritter P, Stewart AL, *et al*. Chronic disease self-management program: 2-year health status and health care utilization outcomes. *Med Care* 2001;39:1217–23.
18 Watters C, Reedy G, Ross A, *et al*. Does interprofessional simulation increase self-efficacy: a comparative study. *BMJ Open* 2015;5:e005472.
19 Pike T, O'Donnell V. The impact of clinical simulation on learner self-efficacy in pre-registration nursing education. *Nurse Educ Today* 2010;30:405–10.
20 Kameg K, Howard VM, Clochesy J, *et al*. The impact of high fidelity human simulation on self-efficacy of communication skills. *Issues Ment Health Nurs* 2010;31:315–23.
21 Lane J, Lane A. Self-efficacy and academic performance. *Soc Behav Pers* 2001;29:687–93.
22 Stajkovic AD, Luthans F, Self-efficacy LF. Self-efficacy and work-related performance: a meta-analysis. *Psychol Bull* 1998;124:240–61.
23 IBM Corporation. IBM SPSS Statistics for Windows, Version 22.0 [program]. Armonk, New York: IBM Corporation. 2013.
24 Gerbing DW, Anderson JC. On the Meaning of within-Factor Correlated Measurement Errors. *J Consum Res* 1984;11:572–80.
25 Brown TA. *Confirmatory factor analysis for applied research*. New York City: Guilford Publications, 2014.
26 Barr H, Low H, Gray R. Introducing interprofessional education*centre for the Advancement of Interprofessional Education (CAIPE)*, 2013.