

The Extended Supervised Learning Event (ESLE): Assessing Nontechnical Skills in Emergency Medicine Trainees in the Workplace



Will Townend, MD, FRCEM*; Alan Gopal, MBBS, MSc; Lynsey Flowerdew, MD, FRCEM;
Amanda Farrow, FRCEM; James Crossley, MBBS, DM

*Corresponding Author. E-mail: william.townend@hey.nhs.uk, Twitter: @willtownend.

Study objective: The contribution of emergency medicine clinicians' nontechnical skills in providing safe, high-quality care in the emergency department (ED) is well known. In 2015, the UK Royal College of Emergency Medicine introduced explicit validated descriptors of nontechnical skills needed to function effectively in the ED. A new nontechnical skills assessment tool that provided a score for 12 domains of nontechnical skills and detailed narrative feedback, the Extended Supervised Learning Event (ESLE), was introduced and was mandated as part of the Royal College of Emergency Medicine assessment schedule. We aim to evaluate the psychometric reliability of the ESLE in its first year of use.

Methods: ESLEs were mandated for all UK emergency medicine trainees in the final 4 years of a 6-year national training program from August 2015. The completed assessments were uploaded to the Royal College of Emergency Medicine e-portfolio. All assessments recorded in the Royal College of Emergency Medicine e-portfolio database between August 2015 and August 2016 were anonymized and analyzed for psychometric reliability, using generalizability theory. Decision analysis was used to model the effect of altering the number of episodes and assessors on reliability.

Results: A total of 1,390 ESLEs were analyzed. The majority (62%) of the variation in nontechnical skills scores was attributable to the trainee's ability. The circumstances of the event (eg, case complexity, workload) accounted for 21% and the stringency or leniency of assessors the remaining 16%. Decision analysis suggests that 3 ESLEs by 2 or more assessors, as currently recommended in the Royal College of Emergency Medicine curriculum, provide an assessment with a reliability coefficient of 0.8.

Conclusion: Board-certified-equivalent emergency medicine supervisors are able to provide reliable assessments of emergency medicine trainees' nontechnical skills in the workplace by using the ESLE. [Ann Emerg Med. 2019;74:670-678.]

Please see page 671 for the Editor's Capsule Summary of this article.

Readers: click on the link to go directly to a survey in which you can provide **feedback** to *Annals* on this particular article.

A **podcast** for this article is available at www.annemergmed.com.

0196-0644/\$-see front matter

Copyright © 2019 by the American College of Emergency Physicians.

<https://doi.org/10.1016/j.annemergmed.2019.05.024>

SEE RELATED EDITORIAL, P. 679.

INTRODUCTION

Background

Providing health care in today's emergency departments (EDs) is complex¹; even a seemingly simple individual act such as prescribing medication calls for a myriad of knowledge and skills beyond simply knowing what to prescribe and how.² Managing a patient safely in an environment of great complexity requires individual skill (ie, situational awareness, the ability to make complex decisions, and adaptability in planning), as well as interpersonal skill (effective communication and leading or being a responsible part of a team).³ These are known collectively as nontechnical skills, which describe the

“human” aspects of performance and overlap heavily with the domain described as human factors.

Nontechnical skills have been considered “soft skills” in health care,⁴ with greater emphasis placed on knowledge and technical skills in training and assessment.^{3,5} This may stem from a lack of clarity over what they are, how to develop them, and how to measure them.⁶ There is a growing realization that human factors underlie a majority of medical errors.⁷ If we are to improve patient safety by better preparing emergency physicians for these aspects of their role, we need to systematically address this through programs of learning and assessment. Given that human error contributes up to 90% of preventable harm in health care,⁸ this can be considered a matter of some urgency.

Editor's Capsule Summary

What is already known on this topic

The Extended Supervised Learning Event (ESLE) is a validated framework for competency assessment of “nontechnical skills” required for effective team leadership in managing complex emergencies.

What question this study addressed

Is the reliability of the ESLE framework adequate to be used effectively with the resources available to standard emergency medicine residency programs?

What this study adds to our knowledge

The authors analyzed reliability of ELSE, using data from 1,390 assessments of 701 trainees in emergency residencies across the United Kingdom. Most of the variation in scores was explained by differences in trainee skills. Two assessors evaluating trainees 3 times a year could achieve acceptable reliability.

How this is relevant to clinical practice

ESLE offers a practical framework for periodic evaluation of multiple resident competencies.

Our previous inability to evaluate such aspects of learning has undoubtedly hampered our ability to build them into specialty training. Recently, however, Flowerdew et al⁹ created an evidence-based behavioral marker system for nontechnical skills developed for use specifically in the ED. Its subsequent evaluation has demonstrated content validity, acceptability, and a test-retest reliability of 0.7 (Spearman's) with a small sample.¹⁰ Such a tool, if it were feasible and reliable on a large scale in routine use, would allow the systematic feedback of nontechnical skills to all emergency medicine trainees. By providing a common taxonomy for such factors, it could potentially support their widespread dissemination and acceptance in the emergency medicine community. The implementation of a systematic approach to assessing and coaching emergency medicine learners in nontechnical skills has yet to be described.

Goals of This Investigation

The Royal College of Emergency Medicine introduced a new assessment strategy for higher training in 2015. It systematically embedded the stated need to assess and support development in nontechnical skills in its assessment schedule. The Extended Supervised Learning Event (ESLE) was introduced and mandated in UK emergency medicine training. It is the instrument developed by Flowerdew et al¹⁰ with a revised rating scale

and is used at intervals through the training year for emergency medicine learners.¹¹ The ESLE form is shown in Appendix E1 (available online at <http://www.annemergmed.com>).

The ESLE offers the opportunity to “grade” observed behavior in the workplace in the domains of nontechnical skill. It also offers the opportunity of detailed narrative feedback. We report the psychometric reliability of the ESLE tool for the graded assessment element of the tool, using generalizability theory. This builds on the work of Flowerdew et al¹¹ in 2 important ways: reliability is evaluated in the intended trainee population (reliability is population dependent), and reliability is assessed with a method that simultaneously estimates the magnitude of measurement error from multiple sources in comparison with the classic test theory approach. Classic test theory evaluates reliability taking only one source of measurement error each time.

MATERIALS AND METHODS

Study Design and Setting

We conducted a generalizability analysis on all ESLE scores recorded for UK emergency medicine trainees in the Royal College of Emergency Medicine e-portfolio from its launch (August 2015) for 12 months.

Emergency medicine specialty training in the United Kingdom typically takes 6 years to complete. It follows 5 years at medical school, usually entered from high school, and 2 years of generic “foundation” training. Trainees are recruited nationally and follow a program that delivers the Royal College of Emergency Medicine curriculum. Training is delivered regionally in training rotations that take trainees through a number of differing EDs. The requirement for assessment, including workplace-based assessment, is the same wherever a trainee is located. These assessments are performed by emergency medicine consultants working in the same ED as the trainees and offer a formalized method for providing feedback on performance. These assessments are considered part of an annual review of competency progression, conducted regionally with external assessors from the training department, that decides on readiness to enter the next year of training.

The first 2 years of emergency medicine training in the United Kingdom are devoted to the care of individual, acutely sick adult patients and include training time in anesthesia, critical care, and internal medicine. Years 3 to 6 are when skills of leadership, overview, and the support of others in the ED are developed and a greater level of responsibility is shouldered, including being the most

senior emergency medicine clinician in the ED, particularly overnight. We therefore introduced the ESLE for all trainees in the United Kingdom in years 3 to 6 in an attempt to assess and support their development.

ESLE was launched for use in all EDs in the 4 nations of the United Kingdom, with supporting material on the Royal College of Emergency Medicine’s Web site describing the objectives of the ESLE for learners and assessors and with material available for promulgation by regional heads of training. Assessors were consultants or equivalent (board-certified equivalent), but there was no requirement that they have particular training in nontechnical skills assessment.

ESLEs were undertaken twice in 6 months for individuals in year 3 of specialty training, 3 times in the year for those in years 4 and 5, and twice in the sixth (final) year. Trainees were informed that the first ESLE should take place in the first 3 months of the training year, the second by 6 months, and the third before the annual review of competency progression, usually held in month 11 of the training year. All of these ESLE assessments were included in analysis.

For ESLE assessments for individuals in years 3 to 5 (inclusive), the assessor was supernumerary and was present solely to observe the trainee in the workplace. The trainee was under the supervision of the consultant or attending clinician on the floor and not the assessor, who acted as an observer. In year 6, the learner was “in charge” of the ED, with the emergency medicine consultant or attending physician on the floor alongside him or her. Each episode was designed to take 3 hours overall, with 1 hour for feedback using the nontechnical skills tool.

Assessors were asked to judge the trainee’s evolving nontechnical skills in 12 domains as rudimentary (A), basic/early (E), sound/higher (H), or consistently high/exemplary (C) (Figure 1). This kind of evolving sophistication judgment aligns with recognizable notions of independence and has been associated with more reliable judgments in other settings.¹² Assessors completed the ESLE form on the Royal College of Emergency Medicine e-portfolio and recorded a rating for all categories of nontechnical skills they believed had been exemplified in the extended episode observed.

'A'	'E'	'H'	'C'
Performance in this aspect expected in core training (Years 1-2 of emergency medicine training)	Performance in this aspect expected at the end of core training or early Higher Specialist Training (Year 3-4 of emergency medicine training)	Performance in this aspect expected in Higher Specialist Training (year 5-6 of emergency medicine training)	Performance in this aspect expected of a consultant (attending emergency physician)

Figure 1. Rating scale for nontechnical skills used in ESLE.

Methods of Measurement

Anonymized data were extracted from the Royal College of Emergency Medicine e-portfolio database. Trainees and assessors were each given unique anonymous identifiers. Training location was not recorded, but year of training was. All complete records in the Royal College of Emergency Medicine database were analyzed.

Outcome Measures

The primary outcome in this study was the reliability of the ESLE for assessing emergency medicine trainees' nontechnical skills, measured with the generalizability coefficient of the ESLE rating tool in this population (the intended population).

In any judgment-based assessment, the score attributed is affected by a range of variables, including the stringency or leniency of the assessor (across all assessments), the difficulty or challenge of the situation (in this case, the shift), the ability of the candidate (across all situations), and the interactions between these variables; for example, whether the way the candidate performed aligned with the subjective preferences of the assessor (there is more than one way of doing things), and whether the candidate's particular aptitudes or experience aligned well with the challenges he or she faced on a particular shift. Classic reliability theory works by conducting separate experiments to test each potential source of variation (interassessor reliability, interchallenge reliability, etc). Generalizability theory takes a different approach by coding all the sources of variation in a larger single "experiment" and applying variance component analysis to estimate the effect of all the primary variables and their interactions. It has largely replaced classic reliability theory in educational assessment validation for 3 main reasons. First, using all the data for all the effect estimates is more powerful and efficient. Second, the effect estimates can be combined by using the simple general equation "reliability=wanted variance/all observed variance (wanted and unwanted)" to produce a simple, meaningful, all-encompassing coefficient of reliability between 0 (completely unreliable) and 1 (completely reliable). Third, those same observed effect estimates can be used to mathematically model the reliability of a hypothetical assessment situation with different numbers of assessors or observations, or with a different observation arrangement.

Our data set is naturalistic; this was an evaluation of ESLE as used nationally rather than a smaller controlled experiment. Therefore, we followed the advice of international experts in the application of generalizability theory to naturalistic data¹² and used

the MINQUE procedure in SPSS,¹³ which produces the best effect estimates for such data sets. Starting with the most sophisticated possible model and then removing variables by reverse stepwise regression until effect estimates were stable and positive, we derived the most sophisticated model supported by the data in line with the authors' recommendations. There was limited trainee-assessor crossing (ie, no examples of 2 assessors judging the same ESLE episode, and not many instances in which 2 or more assessors independently assessed the same 2 or more trainees), so the data set did not support an estimate of that part of assessor subjectivity that can be attributed to preference over particular trainees.

Using the output of the generalizability analysis we performed, we undertook secondary decision analysis to model the effect on reliability of changing the number and arrangement of observations and assessors. In this way, we were able to evaluate our overall approach to nontechnical skills assessment and, if necessary, recommend a refinement to our assessment sampling approach.

Primary Data Analysis

The ordinal categorical ratings A, E, H, and C were assigned a pseudonumeric value (1 to 4, respectively) for analysis. Converting ordinal categories to (pseudo) numbers is common practice in nonphysical measurement (as is assigning a numeric score in the first instance); both uses of numbers make certain assumptions (see "Discussion") but have the advantage of enabling cautious statistical analysis. Aspects recorded as "not observed" were treated as missing data for analysis. However, forms in which no response option had been used, even "not observed," were excluded from the analysis.

We used exploratory factor analysis (principle axis factoring) to examine the relationships between the 12 scored items. In accordance with an eigenvalue threshold of 1.0, there was no evidence of multidimensionality; a single factor accounted for 81.8% of the variance. In other words, all the items appeared to reflect a single underlying factor or trait. (In terms of classic reliability theory, this level of correlation amounts to a Cronbach's α of .98.) Although this in no way diminished the value of the separate elements for feedback purposes, it did allow us to treat the items as psychometrically unidimensional for subsequent analysis. Consequently, we amalgamated the scores across all 12 nontechnical skills domains to give a "global score" for subsequent analysis. For this we took the simple mean of the 12 items, thereby avoiding missing data artifacts.

Generalizability analysis was undertaken with SPSS (version 22).¹³ Generalizability uses variance component analysis (generalizability study) to estimate the contribution that each key assessment variable (or facet) makes to score variance. In this case, the object of measurement was trainee ability (differences between trainees that were consistent across all observations) and facets were the observation episode (episode-to-episode variation) and assessor stringency (differences between assessors that were consistent across all assessments).

Some of these contributions are desirable (eg, trainee ability will influence the score) and some are undesirable (eg, assessor stringency variation should not influence the score). Therefore, the variance estimates can be combined to reflect the overall reliability (R) of the assessment, using the general equation $R = \text{wanted variance} / \text{total variance}$. If only desirable facets influence the score, the reliability will approach 1; if only undesirable facets influence the score, the reliability will approach 0.

Finally, it is possible to mathematically model the effect on reliability or generalizability of varying the number and arrangement of observations and assessors, using a decision study based on the variance estimates. One supervisor's assessment of one performance is unlikely to perfectly predict the result of another supervisor's observation of another single performance; but 100 supervisors' observations of 100 performances are likely to predict the result of a different 100 supervisors' observations of a different 100 performances very well. Exactly how many supervisors or performances are required for reproducibility depends on how much variability each contributes to the score. The decision study simply uses the estimated variance components to model how wide a sample of each key variable is required for the difference between trainees to be "reasonably" reproducible.¹⁴

RESULTS

A total of 1,522 ESLEs were recorded in the Royal College of Emergency Medicine e-portfolio across the United Kingdom between August 2015 and August 2016. One hundred thirty-two had missing data entries and so were removed from analysis. A total of 1,390 ESLEs were entered into subsequent analysis. There were 16,692 data points across all 12 questions in the 1,390 ESLEs included in analysis. Of these, 1,168 (6.9%) were recorded as "not observed." These were assigned the mean nontechnical skills score for that trainee in that ESLE.

In total, 701 trainees were assessed by 750 assessors. The number of trainees and ESLEs undertaken, by year of training, are shown in Table 1. The median number of

Table 1. Number of ESLEs recorded in the Royal College of Emergency Medicine e-portfolio August 2015 to August 2016.

Training Year	No. of Trainees	No. of ESLEs	ESLEs/Trainee
ST3	243	402	Mean 1.65 Median 2 Range 1-4
ST4	427	427	Mean 2.28 Median 2 Range 1-4
ST5	155	332	Mean 2.14 Median 2 Range 1-5
ST6	138	230	Mean 1.66 Median 2 Range 1-4

assessments by each supervisor was 3.00 (interquartile range 2.5 to 3.25). The pseudonumeric nontechnical skills score for each ESLE episode, by year of training, is shown in Figure 2.

The generalizability study results are shown in Table 2. We identified 3 facets that explained all the variance in our analysis, with no residual. Trainee ability had by far the largest influence on ESLE score variation (62%).

The decision study combined all the data to provide the best prediction of generalizability under various assessment conditions. The results are shown in Table 3. These data suggest that 2 assessors providing 1 separate ESLE each can provide a combined result with sufficient reliability for in-training assessment (0.77), and that 3 observations are sufficient for summative assessment (0.80), provided that they are not all from the same supervisor.

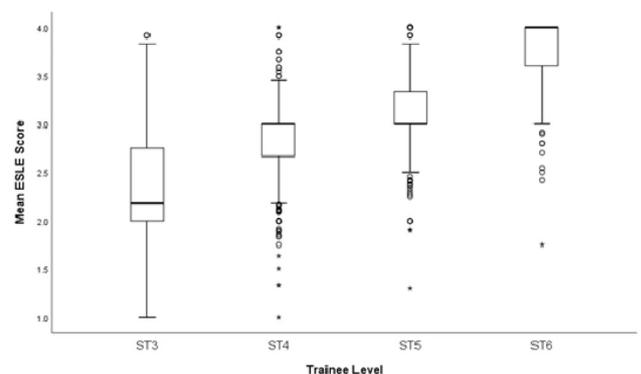


Figure 2. Box plots of amalgamated nontechnical skills score for each ESLE, by year of training. The dark line is the median amalgamated nontechnical skills score for that year of training. The box describes the interquartile range and the bars show 1.5×interquartile range. ST3, Specialty training year 3; ST4, specialty training year 4; ST5, specialty training year 5; ST6, specialty training year 6 (final year of training).

Table 2. Generalizability analysis for ESLE.

Source of Variability	Estimated Variance	
	Component	Variance, %
Trainee	0.295	62
ESLE episode	0.102	21
Assessor	0.077	16
Residual variance	0.00	0

This table shows the result of running generalizability analysis on 1,390 ESLEs recorded. It shows that the variation in amalgamated nontechnical skills scores is related to 3 things: the ESLE episode itself (ie, the variation within them). This might include case mix, time and day, or occurrences during the episode; for example, complex situations or challenging scenarios. The variation because of the trainee is the target for assessment: to what extent the variation in scores is attributable to his or her nontechnical skills. The variation because of the assessor is the fixed effect based on the inclinations of the assessor. These are the elements that are explored in classic methodology when interrater reliability, for example, is assessed. As outlined in the article, the larger the proportion of variance that is attributable to the target of the assessment, the more reliable it is for use.

LIMITATIONS

Conversion of an ordinal categorical scale to continuous numeric data for analysis makes a number of untested assumptions. For example, one cannot empirically test that the intervals between the rating scale points are equivalent at a fundamental level. However, this problem applies to numeric assessment data in the social sciences just as much as ordinal categorical data, and pseudonumeric approximation is a standard operation for analyzing data in the education and social science literature.

In addition to “sophistication” anchors (rudimentary to consistently high), our scale was also aligned to trainee progression (core trainee to consultant) because of the purpose of the assessment. We wanted individual trainees and their supervisors to be able to track progression longitudinally throughout the program in the future. We wanted trainees to know whether they were on track and know whether they were falling behind (and in what areas). The evidence we collected in this study showed that nontechnical skills performance (as measured by ESLE) appears to improve across the training years for individuals following the Royal College of Emergency Medicine training program.

Presenting data by training year as we have does not reflect the potential for a change in scores within year for trainees, and we acknowledge this effect is likely to exist. We consider it likely it will occur in each year of training and is unlikely to greatly affect the difference between years of training we have reported. We handled the data as cross-sectional. It is true that, in a more sophisticated experiment, it might well be appropriate to examine the data for a longitudinal training effect in year, and, if the study were longitudinal across the whole 6-year training cycle, it would be entirely inappropriate to assume that

trainee ability should remain stable precisely because of training effect. We made the assumption that a hypothetical “reliable assessment” (with no session-to-session variation and no assessor stringency or subjectivity variation) would reach the same conclusion each time a trainee was assessed. The natural variation between trainees, and the rate at which they are likely to develop, would mean that our model is more likely to underestimate rather than overestimate the reliability of our tool in the way we analyzed the data.

We must, however, acknowledge another potential artifact. If supervisors simply marked the trainee according to his or her year of training, the scores would increase across the years as an epiphenomenon; however, within each year group, more than half of the trainees were assessed with a progression-based anchor different from their actual year. In future iterations, we plan to ask for a summary view on entrustment for independent practice without reference to year of training, which may offset this potential effect.

The preference of assessors for individual trainees could not be estimated in our study because our data included relatively few assessments of the same trainee by the same trainer. Because assessments were conducted by faculty within the same ED as the trainee and one another, there is also potential for a faculty opinion, favorable or unfavorable, of a trainee to pervade. This is true for any assessment in the workplace. The extended nature of the episode, along with its structure (ie, examples of witnessed behavior used as cues for feedback), may offer an

Table 3. Decision analysis for ESLE modeling the number of ESLEs performed and number of assessors.

Assessors	ESLEs				
	1	2	3	4	5
1	0.62	0.70	0.73	0.74	0.75
2	0.68	0.77	0.80	0.82	0.83
3	0.70	0.79	0.83	0.85	0.87
4	0.71	0.81	0.85	0.85	0.88
5	0.72	0.82	0.86	0.88	0.89

The estimates of variance derived from generalizability analysis have been used to develop a model that describes how the reliability of an assessment tool changes if key elements that drive variance are changed. In our model, we wanted to know the effect on reliability of increasing the number of ESLEs a trainee underwent and the effect of the number of supervisors who made an assessment. Both of these would be expected to lessen the effect on reliability (ie, the effect of the tendency to stringency or leniency of assessors would be diluted if more make assessments on a particular trainee). Decision analysis allows this estimate to be enumerated. This table highlights the reliability of an assessment schedule that includes 3 ESLEs by 2 different assessors. This is what we set out as our requirement for training. It also demonstrates a reliability coefficient of 0.8. This has been used as a threshold for a test that can be used summatively, although that was not our intention here.

Table 4. A comparison of ESLE, CanMEDS educational framework, and Accreditation Council for Graduate Medical Education core competency framework.

ESLE		CanMEDS	ACGME
Maintenance of standards	Subscribes to clinical and safety standards, as well as considering performance targets. Monitors compliance.	Manager Health advocate Professional	Patient care Professionalism System-based practice
Workload management	Manages own and others' workload to avoid both under- and overactivity. Includes prioritizing, delegating, asking for help, and offering assistance.	Manager	System-based practice
Supervision and feedback	Assesses capabilities and identifies knowledge gaps. Provides opportunities for teaching and constructive feedback.	Scholar	Practice-based learning and improvement
Team building	Provides motivation and support for the team. Appears friendly and approachable.	Collaborator	Interpersonal and communication Professionalism
Quality of communication	Gives verbal and written information concisely and effectively. Listens, acknowledges receipt of information, and clarifies when necessary.	Communicator	Interpersonal and communication
Authority and assertiveness	Behaves in an appropriately forceful manner and speaks up when necessary. Resolves conflict effectively and remains calm when under pressure.	Collaborator Professional	Interpersonal and communication Professionalism
Option generation	Uses all resources (written and verbal) to gather information and generate appropriate options for a given problem or task. Involves team members in the decisionmaking process.	Medical expert Collaborator	Medical knowledge
Selecting and communicating options	Considers risks of various options and discusses this with the team. Involves clearly stating decisions and explaining reasons, if necessary.	Medical expert Communicator	Patient care Medical knowledge Interpersonal and communication
Outcome review	Once a decision has been made, reviews suitability in light of new information or change in circumstances and considers new options. Confirms tasks have been done.	Manager	Patient care Interpersonal and communication
Gathering information	Surveys the environment to pick up cues that may need action, as well as requesting reports from others.	Manager	System-based practice
Anticipating	Anticipates potential issues such as staffing or cubicle availability in the department and discusses contingencies.	Medical expert Manager	Patient care System-based practice
Updating the team	Cross-checks information to ensure it is reliable. Communicates situation to keep team informed rather than just expecting action.	Communicator Collaborator	Interpersonal and communication System-based practice

ACGME, Accreditation Council for Graduate Medical Education.

opportunity to gain a more valid impression of capability. Our results, in comparison with other modes of assessment in the workplace, support that interpretation.

DISCUSSION

We report here the implementation of ESLE nationally to all UK emergency medicine trainees in the latter part of a 6-year training program. The generalizability analysis we performed estimated that 62% of the variation in mean nontechnical skills scores from ESLE was related to the trainee's nontechnical skills. When using these data to perform a decision analysis, we found that 3 ESLEs were required to achieve high levels of reliability. This compares

very favorably with other workplace-based assessments, in which typically 8 to 20 observations may be required. More importantly for our purposes, ESLE is one of the first instruments specifically designed to assess nontechnical skills.

The high reliability is likely to reflect, at least in part, the duration of the observation episode, which allows extended scrutiny and judgment on the part of the assessor. Also, we used a scale aligned to notions of evolving sophistication and independence in training that we believed would be understandable to assessors. Weller et al¹⁵ showed this to be an effective approach to improving reliability. The ESLE is also built on a nontechnical skills behavioral marker system

with established validity in the emergency medicine workplace.

The ESLE evolved from a nontechnical skills framework that was based on published emergency medicine safety literature⁹ and UK curricula.¹⁰ Comparison with other international educational frameworks used to frame nontechnical skills demonstrates considerable commonality (Table 4) and supports the construct validity of this assessment tool.^{16,17} It also suggests our instrument might have a wider role in health care settings in which nontechnical skills are considered key.

It has been suggested that detailed nontechnical skills training is required before such complex attributes can be evaluated in the workplace.¹⁸ Our data suggest otherwise. Using only the ESLE instrument and relatively limited guidance, our faculty were able to make generalizable assessments of “high-level” attributes in the ED. The launch of ESLE included training materials made available to regional heads of training and a description on the Royal College of Emergency Medicine Web site. It is likely, therefore, that formal training in nontechnical skills was variable among the assessors in this data set. Given that a comprehensive program of training may represent a significant logistic challenge, our findings are particularly reassuring for the future use of this tool. We believe our data suggest that board-certified-equivalent emergency medicine clinicians have valuable insights into nontechnical skills that can be captured and shared. It is possible that with experience and training, the reliability of nontechnical skills assessment will also improve.

In summary, we report that ESLE was launched nationally as part of UK emergency medicine training in 2015. It appears to be reliable for assessing emergency medicine trainees from the start of higher training to independent practice. We believe this is the first such tool to be introduced for universal use, and indeed our WPBA schedule in higher training depends heavily on it. The ESLE also provides a taxonomy for reflecting and addressing development needs and providing feedback in the 12 chosen areas of nontechnical skills. In this study we did not specifically report feasibility or any measure of educational influence, but clearly learning and feedback are integral to the design of this instrument, and we plan to report on this in a future study.

It is our belief that, by improving the monitoring and feedback of nontechnical skills, our trainees will be better prepared for key elements of their future roles.

This study was performed on behalf of the Royal College of Emergency Medicine by members of its curriculum development group. We wish to acknowledge the commitment

and professionalism of trainers and trainees in UK Emergency Medicine in piloting this new workplace-based assessment prior to national launch. We also wish to acknowledge the adaptability and ability of UK emergency medicine trainers and trainees to embrace innovation.

Supervising editor: Peter C. Wyer, MD. Specific detailed information about possible conflict of interest for individual editors is available at <https://www.annemergmed.com/editors>.

Author affiliations: From Emergency Department, Hull Royal Infirmary (Townend) and Hull Institute of Learning and Simulation (Gopal), Hull and East Yorkshire Hospitals NHS Trust, Hull, East Yorkshire, United Kingdom; Emergency Department, Frimley Park Hospital, Frimley Health NHS Foundation Trust, Camberley, Surrey, United Kingdom (Flowerdew); Emergency Department, Princess of Wales Hospital, Bridgend, Mid Glamorgan, Wales, United Kingdom (Farrow); and The Medical School, University of Sheffield, Sheffield, South Yorkshire, United Kingdom (Crossley).

Author contributions: WT, LF, and AF developed the tool and took it through a pilot phase. WT and JC designed the study. WT and AG collated data. WT, AG, and JC analyzed the data and conducted the generalizability analyses. WT wrote the first draft of the article and all authors contributed significantly to its editing. WT takes responsibility for the paper as a whole.

All authors attest to meeting the four [ICMJE.org](http://www.icmje.org) authorship criteria: (1) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND (2) Drafting the work or revising it critically for important intellectual content; AND (3) Final approval of the version to be published; AND (4) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding and support: By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see www.icmje.org). Dr Townend has received payment to work for the Royal College of Emergency Medicine on curriculum development, but was not paid in relation to this current study. Dr. Crossley reports having undertaken paid educational consultancy work for the Royal College of Emergency Medicine but not in connection with the current study.

Publication dates: Received for publication January 18, 2018. Revisions received June 20, 2018; October 16, 2018, and May 1, 2019. Accepted for publication May 10, 2019.

REFERENCES

1. Chisholm C, Collison E, Nelson D, et al. Emergency department workplace interruptions: are emergency physicians “interrupt-driven” and “multitasking”? *Acad Emerg Med*. 2000;7:1239-1243.
2. Tully M, Ashcroft D, Dornan T, et al. The causes of and factors associated with prescribing errors in hospital inpatients: a systemic review. *Drug Saf*. 2009;10:819-836.
3. Fletcher G, McGeorge P, Fiin R, et al. The role of non-technical skills in anaesthesia: a review of current literature. *Br J Anaesth*. 2002;88:418-429.

4. Cooper S, Endacott R, Cant R. Measuring non-technical skills in medical emergency care: a review of assessment measures. *Open Access Emerg Med.* 2010;2:7-16.
5. Yule S, Flin R, Paterson-Brown S, et al. Non-technical skills for surgeons in the operating room: a review of the literature. *Surgery.* 2006;139:140-149.
6. Flowerdew L, Brown R, Vincent C, et al. Identifying nontechnical skills associated with safety in the emergency department: a scoping review of the literature. *Ann Emerg Med.* 2012;59:386-394.
7. Kohn LT, Corrigan JM, Donaldson MS; Institute of Medicine. *To Err Is Human: Building a Safer Health System.* Washington, DC: National Academy Press; 2000.
8. Corrigan J, Donaldson M, Kohn L; Committee on Quality of Health Care; Institute of Medicine. *To Err Is Human: Building a Safer Health Care System.* Washington, DC: National Academy Press; 1999.
9. Flowerdew L, Brown R, Vincent C, et al. Development and validation of a tool to assess emergency physicians' nontechnical skills. *Ann Emerg Med.* 2012;59:376-385.
10. Flowerdew L, et al. A multicentre observational study to evaluate a new tool to assess emergency physicians' non-technical skills. *Emerg Med J.* 2013;30:437-443.
11. Royal College of Emergency Medicine. 2015 RCEM WPBA schedule: curriculum—appendix 1. Available at: [https://www.rcem.ac.uk/docs/Training/2015%20Curriculum%20Appendix%201%20\(July%202016%20Update\).pdf](https://www.rcem.ac.uk/docs/Training/2015%20Curriculum%20Appendix%201%20(July%202016%20Update).pdf). Accessed June 19, 2018.
12. Crossley J, Russell J, Jolly B, et al. "I'm pickin' up good regressions": the governance of generalisability analyses. *Med Educ.* 2007;41:926-934.
13. IBM Corp. *IBM SPSS Statistics for Windows, Version 22.0.* Armonk, NY: IBM Corp; 2013.
14. Crossley J, Davies H, Humphris G, et al. Generalisability: a key to unlock professional assessment. *Med Educ.* 2002;36:972-978.
15. Weller JM, Misur M, Nicolson S, et al. Can I leave the theatre? a key to more reliable workplace-based assessment. *Br J Anaesth.* 2014;112:1083-1091.
16. Accreditation Council for Graduate Medical Education. Accreditation Council for Graduate Medical Education (ACGME) core competencies. Available at: http://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRs_2017-07-01.pdf. Accessed June 1, 2018.
17. Royal College of Physicians and Surgeons of Canada. CanMEDS 2015 physician competency framework. Available at: http://canmeds.royalcollege.ca/uploads/en/framework/CanMEDS%202015%20Framework_EN_Reduced.pdf. Accessed June 1, 2018.
18. Graham J, Hocking G, Giles E. Anaesthesia non-technical skills: can anaesthetists be trained to reliably use this behavioural marker system in 1 day? *Br J Anaesth.* 2010;104:440-445.

IMAGES IN EMERGENCY MEDICINE

(continued from p. 633)

DIAGNOSIS:

Anomalous right coronary artery from the pulmonary artery. Pulmonary artery short-axis view showed one coronary artery originating from the pulmonary artery (Figure 1), and the low-velocity diastolic flow was indicative of septal coronary collaterals (Figure 2), making an anomalous coronary artery the most likely diagnosis.¹ CT angiography (Figure 3) confirmed the origin of her right coronary artery from the main pulmonary artery, with an aneurysm in its proximal part and collaterals between the left and right coronary arteries. The patient underwent reimplantation surgery and was discharged home with medical therapy.

Anomalous right coronary artery from the pulmonary artery is a rare congenital heart disease affecting 0.002% of the population. It has not been considered to be a lethal defect in infancy or childhood, although cases of sudden death have been reported.² Continuous murmur, angina, dyspnea, and cyanosis are frequent presentations of anomalous right coronary artery from the pulmonary artery. Aortopulmonary window and tetralogy of Fallot are the most common associated cardiac lesions.^{2,3} Visualizing the anomalous origin of the coronary artery can be challenging by transthoracic echocardiography. All cases diagnosed by transthoracic echocardiography are confirmed by angiography.⁴ The most common surgical procedure is reimplantation, whereas surgical ligation of the anomalous right coronary artery is performed in a few cases.²

Author affiliations: From the Department of Cardiovascular Surgery, West China Hospital, Sichuan University, Chengdu, Sichuan, People's Republic of China.

REFERENCES

1. Frommelt MA, Miller E, Williamson J, et al. Detection of septal coronary collaterals by color flow Doppler mapping is a marker for anomalous origin of a coronary artery from the pulmonary artery. *J Am Soc Echocardiogr.* 2002;15:259-263.
2. Williams IA, Gersony WM, Hellenbrand WE. Anomalous right coronary artery arising from the pulmonary artery: a report of 7 cases and a review of the literature. *Am Heart J.* 2006;152: 1004.e9-17.
3. Bortolotti D, Casarotto D, Betti D. Anomalous origin of the right coronary artery from the main pulmonary artery. *Eur J Cardiol.* 1978;7:451-455.
4. Donaldson RM, Raphael M, Radley-Smith R, et al. Angiographic diagnosis of anomalous origin of the right coronary artery from the pulmonary artery. *Br J Radiol.* 1983;56:17-19.