ORIGINAL RESEARCH

a Open Access

Do supervised learning events reflect clinical competency and support "trainee in difficulty" identification? Physician trainees' perceptions

Andrew Tomkins[1,3], Cathy Sherratt[4*], Mumtaz Patel[2,3*]

[1]Faculty of Health and Social Care, Edge Hill University, Ormskirk, UK

[2]Health Education England, Manchester, UK

[3]Manchester University Hospitals NHS Foundation Trust, Manchester, UK

[4]Faculty of Health, Social Care and Medicine, Edge Hill University, Ormskirk, UK

[*]Both authors contributed equally in the lead author role.

ABSTRACT

Objective: Training and assessment of postgraduate medical trainees has undergone a process of standardization in recent years by using Workplace-Based Assessments (WPBAs). WPBAs play a pivotal role in assessing competency and ensuring satisfactory training progress. From 2012 onwards, traditional WPBAs in the UK were replaced by Supervised Learning Events (SLEs) that include substantial formative feedback. SLE use is encouraged in the identification and monitoring of training difficulties. Trainees' perceptions of their value in identifying training difficulties and assessment of clinical competency are yet to be explored.

Methods: A mixed-methods study adopting Grounded Theory methods was conducted with Higher Specialty Trainees across three medical disciplines; individuals with several years of postgraduate experience. Participants completed an online questionnaire utilizing both qualitative and quantitative questions ($n = 25$). Subsequently, two focus groups were conducted to explore perceptions of the assessment process ($n = 14$). Grounded Theory methods were used to develop codes for the qualitative data, with quantitative responses recorded using Likert rating scales.

Results: Multi-rater assessments were rated the highest at assessing clinical competency, with directly observed assessments rated the lowest.

Five main themes emerged from the data:

1. Trainees attempted to present their "best-self": tension was identified between formative and summative aspects of assessments.
2. Assessment process mistrust: concerns regarding the permanency of recording suboptimal performance impaired assessment use.
3. Cultural shift of feedback provision: an enhanced feedback culture was identified, with assessments acting as a "springboard" for knowledge development.
4. Assessor dependence: pivotal role the assessor plays in training difficulty identification.
5. Task-specific nature: narrow remit led to assessments' limited ability to capture trainee performance.

Conclusions: Physician trainees associate SLE introduction with enhanced identification of training difficulties through an improved feedback culture. Threats to optimal SLE use include fear of repercussions of negative outcomes and trainees masking weaknesses.

Contact Andrew Tomkins ✉ andrewtomkins@nhs.net ▭ Manchester University Hospitals NHS Foundation Trust, Manchester, UK.

Andrew Tomkins, Cathy Sherratt, Mumtaz Patel

Introduction

In postgraduate medical training curricula, the use of Workplace-Based Assessments (WPBAs) has played an increasingly important role in supporting learning and assessing competency [1]. Designed to assess clinical activities encountered by doctors on a daily basis, they include a range of validated single- and multi-rater assessment tools that establish performance and provide feedback to enhance future practice [2]. Possessing high validity, their value in assessing trainee performance is well established when combined with traditional "high-stakes" examinations [3]. WPBA use in postgraduate medical training is reported in the United Kingdom, Europe, USA, Canada, and, more recently, parts of Asia [4–7].

Despite their widespread use, concerns have been raised regarding WPBA value; with reports of both trainees and trainers perceiving them as onerous and "tick-box" exercises to meet training curriculum requirements [3,8,9]. Accordingly, "Supervised Learning Events" (SLEs) were introduced across training programme curricula in the UK to help address these limitations [10]. While based on pre-existing WPBA tools, the emphasis was placed on encouraging the provision of timely feedback, enhancing the trainee–trainer interaction through more structured dialog, a renewed focus on both formative and summative aspects of the assessment process and greater engagement in the educational process [11]. Rating scales were replaced with free-text responses to encourage feedback and reflection. Additional multi-rater assessments including the multiple-consultant report (MCR) were introduced. Initially piloted in 2012, their use became widespread across all disciplines in 2014 [12]. Table 1 compares traditional WPBAs with newer SLEs.

In the United Kingdom, between 2% and 6% of the medical workforce are anticipated to demonstrate difficulty sufficient enough to raise concerns regarding their competency [13,14]. Defined as an individual "who demonstrates a significant enough problem that requires intervention by someone of authority," trainees in difficulty (TiD) may adversely impact patient care, their own training, and the wider healthcare team [15]. While multiple characteristics and presentations of TiD have been identified, the most common relate to those involving knowledge, skills, or attitudes [16]. Both national and international guidance have been developed to help support Clinical and Educational Supervisors in the identification and management of TiD [17,18]. This highlights the importance of early identification and intervention to maintain patient, trainee, and colleague safety and to prevent the issues becoming intractable. Guidance encourages the use of SLEs to help document and monitor knowledge and skill deficiencies once TiD status has been established.

A literature review of trainees' perceptions of SLE value in establishing clinical competency and assessing TiD status identified a paucity of existing published reports, particularly, following SLE introduction. While a limited number of publications have explored the value of assessments as tools for identifying training difficulties, comparisons of pre- and post-SLE introduction have yet to be made [14,19,20]. Multi-rater assessments appear most beneficial in identifying factors relating to training difficulties [14]. Limited research involving higher specialty trainee [Specialty Trainee Year 3 + (ST3)] participants exists and no studies specifically focus on physician trainees; a cohort which is likely to both have had experience of traditional WPBA and newer SLEs. Consequently, the aim of this study is to address the following questions:

- How do physician trainees perceive newer SLE tools in comparison to traditional WPBAs in assessing clinical competency?
- What is the perceived value of SLEs in identifying trainees in difficulty and how could they be improved?
- What are physician trainees' experiences of assessing colleagues and do these accurately reflect the competency of the individual?

Methods

This dual-phase mixed-methods study which adopted Grounded Theory principles [21] involved an online questionnaire of higher specialty trainees (ST3–ST7) across three physicianly specialties which dual train with General Internal Medicine (Geriatric Medicine, Diabetes and Endocrinology and Genitourinary Medicine). Grounded Theory is a research approach whereby analysis and generation of theory arise out of and are "grounded in" the data itself [21], rather than relying on a hypothesis or any preconceptions on the part of the researchers. Analysis of qualitative data commences with simple descriptions, arising naturally out of the data, which are then refined into analytic themes, thus allowing the lived experience of participants

Table 1. Traditional workplace-based assessments versus supervised learning events in the UK postgraduate medical curriculum (Adapted from Collins Report 2010 [8] and The Joint Royal College of Physicians recommendations for Specialty Trainee Assessment 2014 [12].

Traditional workplace-based assessment	Supervised learning events
Focus on summative assessment	Focus on formative assessment
Rating scale of competency forms a major component of the assessment form	Free text boxes form a major component of the assessment form
Large quantity of assessments to be completed throughout the training year	Fewer assessments, with an emphasis placed on trainee reflection and action planning
Poorer trainee/trainer engagement, with forms often being completed retrospectively	Focus on constructive feedback
	Encourages highlighting of achievements
	Immediate feedback provided with suggestions for areas for development

to be fully acknowledged [22]. The questionnaire established personal experiences of SLEs and perceived value of identification of training difficulties. Subsequently, two focus groups were conducted to explore and triangulate findings from the questionnaire. Both Health Education England (HEE) Research Governance approval (dated 11th May 2017) and Edge Hill University Research Ethics approval (reference FOHS200; dated 21 March 2018) were obtained.

Participant identification

Potential participants were recruited across three medical specialties in an HEE training region in the North West of England. Purposive sampling was conducted across physicianly disciplines to maximize participant homogeneity. Additionally, the participants had experiences using both the traditional WPBAs and the newer SLEs, as well as having experience completing them for other colleagues, both at a junior and senior level. Approval for the respective medical specialties was granted by the Training Programme Director (TPD) prior to approaching potential participants.

Participant contact

Initial contact was made via email by the primary researcher, which included a Participant Information Sheet (PIS) and consent form. Time and opportunity were provided for potential participants to contact the primary researcher for any required clarification.

Phase 1: Questionnaire

An anonymous questionnaire was devised, hosted via a secure survey website [23]. Questions focused on headings from both national and regional guidance, which lists behavioral markers of TiD status.

Qualitative (free text) response and quantitative (Likert rating scale) questions were utilized. The survey was distributed via work email addresses to 89 medical specialty trainees (25 respondents) with PIS attached. The survey was available for 6 weeks from the 22nd June to 3rd August 2017, with a reminder email sent after 4 weeks. Questions were based around four themes: trainees' personal experiences of SLEs, the value of SLEs in identifying clinical competency and improvements, the value in identifying trainees in difficulty, and experiences of completing assessments for other colleagues. Anonymity was preserved by the absence of any requests for identifiable information and suitable anonymization of any direct quotations. The questionnaire template can be found in Appendix A.

Phase 2: Focus groups

Potential participants in the same medical specialties and the same cohort as phase one were invited to participate in the second phase. It was not necessary for focus group participants to have completed the phase one, nor was this information requested in order to preserve the anonymity of questionnaire participants. Potential participants were initially contacted via email 2 weeks prior to the focus group dates, with the PIS and consent form attached. A suitable training day for conducting the research was agreed upon by the TPD and a trainee representative. Focus groups were conducted following regional specialty teaching days, both for participant convenience and to maximize participation. The study was introduced by the primary researcher, which allowed potential concerns to be addressed. Individuals were under no obligation to participate and were able to withdraw their consent. Confidentiality was ensured through: informing

participants during the focus group pre-amble that no references to individual hospital sites or people should be made and removal of identifiable information at the point of transcription. Data from the questionnaire informed the focus group schedule. To ensure the accuracy of data interpretation, audio recording devices were used during the focus groups. The participants were made aware of this, both on the PIS and prior to conducting the focus group and they provided written consent agreeing to this. The groups lasted approximately 45 minutes each. Data were transcribed verbatim and anonymized at the point of transcription. Data were covered by the Data Protection Act (2018) [24]. A copy of the Focus Group Schedule can be found in Appendix B.

Data analysis

Data analysis commenced immediately following the collection of questionnaire data, in line with Grounded Theory principles [25]. This enabled theoretical sampling whereby simultaneous data collection, coding, and analysis occurred. This allowed the initial data to be used to inform subsequent steps taken that informed the development of the questions in the second phase. Ensuring line by line coding and memo-writing was used throughout the process, the data were analyzed until theoretical saturation was achieved. This occurred during the analysis of the second focus group transcript, making a third focus group unnecessary. These ensured concepts were well-developed and provided an end-point for data collection [22]. The primary researcher conducted the initial analysis which was verified by all members of the research team and a consensus reached regarding the emergent themes. Using the Grounded Theory approach allowed nuance to arise from the data, with different aspects being explored than had initially been considered.

Results

Twenty-five participants completed the questionnaire (28% response rate); 14 individuals participated in the focus groups (focus group one: six participants; focus group two: eight participants). All participants had experience of both traditional WPBA assessments and the newer SLEs, as well as having experience completing these for other colleagues. Each participant was assigned a numerical code (P) with the results presented below and in the diagrams. Quantitative results in the form of Likert

rating scales regarding individual assessments were identified. Analysis of qualitative data identified five main themes (summarized in Table 2). Example quotations that support the themes are included in Table 3.

Main themes

Theme 1: Trainee drive to present the "best-self" during the assessment process

Participants reported both a personal drive to ensure only positive assessment outcomes were recorded on the e-portfolio and also when witnessing the same behavior in other trainees through their role as an assessor. Tension between receiving constructive feedback, (identified as valuable for enhancing future performance) and the fear of adverse outcomes from recording less than satisfactory assessments existed. Participants also reported witnessing trainees "mask" areas of weakness through a preferential selection of assessors who were deemed lenient, or selection of a topic they felt more confident in, while avoiding complex topics.

Suggestions made by participants to minimize the potential adverse impacts of this included placing the responsibility for monitoring the e-portfolio at regular intervals on the Educational Supervisor, to ensure that a sufficient selection of assessors and breadth of topics are being covered. Additionally, placing the onus on the assessor to select cases for discussion was also suggested, to help limit the possibility of knowledge deficiencies.

Theme 2: Mistrust of the assessment process

Participants reported a wariness of the assessment process, with concerns relating to the permanency of documentation on the SLEs and the potential repercussions on their own, or other colleagues' training progression. A tendency to self-edit and ensure only positive comments are recorded was reported following concerns that information on the e-portfolio may be used against trainees in certain circumstances. A reluctance to engage fully with the assessment process was subsequently reported. The importance of being prepared to discuss comments which are documented in the assessment to the trainee via a face-to-face dialog was cautioned. Participants reported confidentiality breaches that had occurred despite feedback having been given under seemingly anonymous circumstances; for example, during the completion of a multi-source feedback (MSF).

Table 2. Themes and concepts identified from questionnaire and focus groups.

Present "Best-Self" Image	Mistrust of Assessment Process	Enhanced Feedback culture	Role of Assessor	Narrow scope of SLEs
Concepts	**Concepts**	**Concepts**	**Concepts**	**Concepts**
• Failure to view SLEs as formative as well as summative • Tension between learning from constructive feedback and concerns regarding adverse outcomes • Trainees 'mask' areas of weakness • Preferential selection of assessors and assessment topics • Fear of documenting less than satisfactory assessments	• Concerns regarding giving or receiving honest feedback and the adverse impact it may have on training • Permanency of documentation on e-portfolio and repercussions • Breaches of confidentiality • Self-editing reflections to ensure no negative repercussions occur • Inherent bias of the assessment process	• More meaningful feedback provision • SLEs break "norm" of receiving negative/critical feedback only • Enhanced individualized feedback • Act as springboard for learning and development • Challenges of the provision of honest feedback and potential for negative implications	• Benefits of formalized training improving standardization across supervisors • Impact of individual assessors' views and opinions of SLEs impacts their value • Pivotal role assessors play in Trainee in Difficulty identification • Need for face-to-face completion of SLEs (versus remote completion) • Power-shift when trainee becomes assessor and the impact of this	• Tension between formative and summative components • SLE requires "breadth" of assessment; too task-specific • Challenges capturing difficulties in e-portfolio • Lack of cohesion across e-portfolio • Issues regarding accessibility
Suggestions for Improvement	**Suggestions for Improvement**	**Suggestions for Improvement**	**Suggestions for Improvement**	**Suggestions for Improvement**
• More regular review of progress by Educational Supervisor • Increase the breadth of assessment • Onus of assessment selection placed on assessor	• Increase time allocated to the assessment process • Ensure timely discussion takes place regarding training concerns • Ensuring clear guidelines and processes are in place	• Ensure comments on SLEs regarding future learning points are mandatory • Enhanced face-to-face dialog when completing assessments	• More dedicated time spent completing assessments in assessor's presence • Ongoing training and revalidation of assessors • Educational supervisor plays a more active role in monitoring progress	• Increase the breadth of assessment • More onus placed on multi-rater assessments • Ensure assessments specifically address clinical competency • Provision of an option for confidential feedback to be sent to the trainee's Educational Supervisor

Suggestions for improvement included increasing the time allocated to the assessment process which would allow deeper working relationships to establish as well as streamlining the assessment process to allow the improved depth of completion. Additionally, ensuring trainees are made aware of training concerns at an early stage was encouraged, rather than such concerns being raised towards the end of their training year. These suggestions are in keeping with international guidance, which stress the critical importance of early identification and intervention in TiD management [18].

Theme 3: Cultural shifts in the provision of feedback

Participants reported an enhanced feedback culture following the SLE introduction. This was viewed as being beneficial both for learning and in the identification of training difficulties. Positive changes included structural revisions to the assessment forms, with a movement from tick-box

Table 3. Example quotations of key themes.

Theme	Example quotation
Present "Best-Self" Image	*"You picked it probably because you felt comfortable doing it rather than something you feel weak on"* (P27)
	"After the assessment, they've said please don't, I don't want you to do that, I'm not sending you a form and it wasn't bad, it was appropriate for the level, and that's when improving, that's when it's a positive thing." (P 28)
Mistrust of Assessment Process	*"I have found myself writing something and then taking it back and rephrasing it because I don't want to adversely affect someone's outcome."* (P31)
	"I've filled in feedback for people and it has gone horribly wrong where the confidentiality has been broken down, it was a long time ago now, but I put an honest review of that person and I should have said it to their face." (P 29)
Enhanced Feedback culture	*"It's probably the only time someone will actually say this is what you did well...I think it's actually quite good to get some positive feedback."* (P36)
	"There's lots of training for how to give [feedback] but there's not much for how to receive it. I think that is hard because it depends, some people are going to be receptive to it and some people just aren't." (P 29)
Role of Assessor	*"if it's done with the person next to you, that's usually what makes it a more worthwhile experience"* (P33)
	"...it is just that chance to have a sit down chat with a consultant that you actually wouldn't otherwise have the time to do." (P 28)
Narrow scope of SLEs	*"They might do well in a specific assessment but you have issues with their overall performance or specific aspects that are not well captured in WPBA assessments"* (P2)
	"They [MSFs and MCRs] are potentially used as evidence from a breadth of assessors with regards to helping the supervisor make an assessment of training difficulties." (P 14)

to free-text responses. Participants also reported the benefits of receiving individualized feedback on their performance and that SLEs "break the norm" of receiving negative feedback in the workplace. SLEs were also reported as being invaluable as a "springboard" for learning, generating areas for further education, and development as well as acting as a benchmark from which new knowledge can be assessed. Several participants, however, reported urging caution regarding the provision of honest feedback. This included a fear of repercussions, the potential for comments to be misinterpreted when documented in an assessment form, and the need to continue to work alongside colleagues following either giving or receiving potentially critical feedback. Additionally, it was felt that an expectation existed among trainees that assessors should give positive feedback only.

Further suggestions for improvement included making the "comments" and "future learning points" sections mandatory and ensuring sufficient time was allocated to assessments to ensure more in-depth feedback was provided.

Theme 4: Pivotal role of the assessor in competency assessment and TiD identification

Participants highlighted the critical role the assessor holds in ensuring competency is assessed and

training difficulties are identified. There was a consensus on the benefits of face-to-face versus remote completion of SLEs. This included having dedicated time with consultant colleagues which otherwise proves challenging to organize.

While it was acknowledged that SLE completion should be trainee-driven, participants reported the benefits of consultants in having dedicated training on how to complete assessments. The assessors' views on SLE value were also felt to be strongly influential in the educational experience of the trainee. When asked how SLEs could be improved to help identify TiD, suggestions included a greater onus placed on the assessors' role and more dedicated time spent on completing assessments, as well as the Educational Supervisor to oversee the selection of assessors to help prevent masking of weaknesses. Additionally, ongoing revalidation of supervisors was suggested to ensure a more standardized approach towards assessment and provision of feedback.

Theme 5: Narrow scope of SLEs may limit TiD identification

Several participants criticized SLE forms for being too task-specific and having a narrow focus. It was acknowledged that TiD may be able to complete certain specific tasks; however, may lack overall

clinical competency or be experiencing training difficulties. This may lead to challenges in such difficulties being captured on the e-portfolio.

Suggestions to navigate the task-specific nature of certain SLEs included questions specifically addressing clinical competency and the option for assessors to confidentially contact the trainees' supervisor to raise any concerns identified. Additionally, there were concerns that identification of training difficulties was often made towards the end of the training year, for example, via the MSF, and thus the opportunity for early intervention has been missed. Other participants felt the inclusion of additional multi-rater assessments such as the MCR enhanced the potential for TiDs to be identified.

Value of individual Supervised Learning Events in assessing clinical competency

Using a Likert scale, questionnaire participants were asked to rate the value of individual assessments when an assessment of competency was being made (Fig. 1). The multi-rater assessments including MSF and MCR were rated the highest as being either very or somewhat useful in assessing competency, whereas the Acute Care Assessment Tool (ACAT) and Clinical Evaluation Exercise (mini-CEX) were rated the least useful.
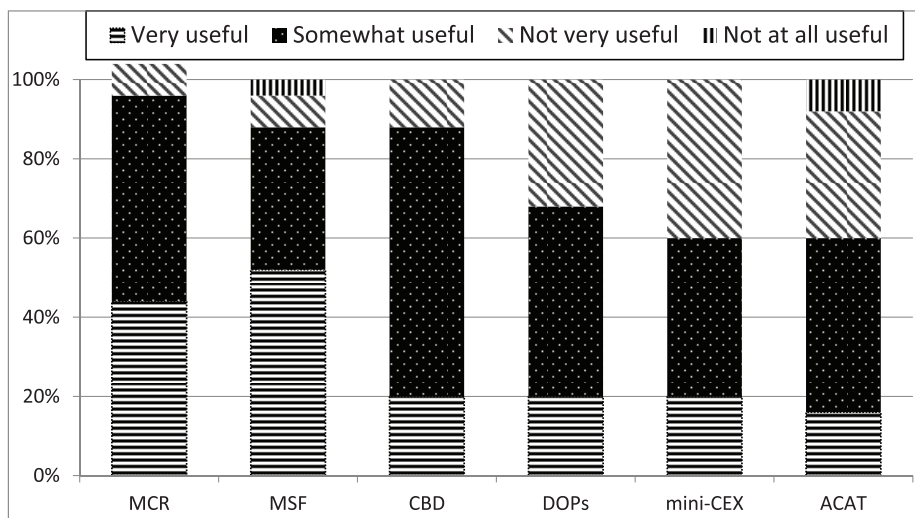
Discussion

The need for an appreciation of the complexities surrounding workplace-based learning and assessment to ensure the maximal benefit is achieved is evident [26]. To the authors' knowledge, this is the first report providing an insight into physicianly higher specialty trainees' perception of the value of SLEs in identifying training difficulties and assessing competency. Previous research predates the introduction of SLEs or focuses on assessor perspectives and acceptability of the assessments [27–29]. Findings build upon the established intricacies involved in using the same tool for multiple purposes, including an indicator for training deficiencies, an educational experience, and an assessment of knowledge [26,30].

Improvements following SLE introduction

These findings demonstrate a tentative improvement in physician trainees' perceptions of SLE value in identifying TiD through an enhanced feedback culture. This is well supported as being critical in early identification of training deficiency [31]. Findings provide some reassurance, as previous research demonstrates polarizing views on feedback quality and educational value of WPBAs [32,33]. An enhanced focus on multi-rater assessments, including the introduction of the MCR, was felt to be beneficial in overall competency assessment and provision of high-quality feedback. This contrasts with research which explored Core Medical Trainees' (UK trainees on the first stage of physicianly training) perceptions of the MCR, reporting participants felt feedback was often ineffective and the assessments duplicated other forms [34]. This may suggest that trainees at a more advanced level

Figure 1. Trainees' rating of individual SLEs in assessing clinical competency. MCR = Multiple Consultant Report, MSF = Multi-Source Feedback, CBD = Case-based Discussion, DOPS = Directly Observation of Procedural Skills, Mini-CEX= Mini Clinical Evaluation Exercise, ACAT = Acute Care Assessment Tool.

may be more receptive to feedback from consultant colleagues than their junior counterparts. The drive to ensure all Educational and Clinical Supervisors receive formalized training was also valued by participants, many of whom were undertaking such training themselves and implementing it in their own practice.

Supervised learning event limitations

Undoubtedly, ongoing issues persist that limit SLEs reaching their full potential in assessing competency and identification of trainees in difficulty. Participants identified inherent deficiencies in the assessment process, including the potential for trainees to conceal areas of weakness and present only positive encounters in their e-portfolios. These findings are supported by research reporting that surgical trainee participants sought positive feedback only when WPBAs were viewed as being summative, despite knowing that critical feedback enhanced performance [33]. Additionally, the research identified trainee-selected assessors to score less harshly than Clinical Supervisors in specialty trainees' MSF assessments [35]. Strategies to navigate this must take into account an increasing wariness reported by trainees to document adverse assessment outcomes [36].

Strengths and limitations

Strengths of this study included a selection of participants who have had experience of both the traditional and new assessment forms, as well as with the completion of assessments for other colleagues, allowing a range of experiences to be captured. The questionnaire provided the advantage of a reduced level of "observer bias," since the anonymity offered afforded participants the opportunity to respond freely. Theoretical sampling allowed phase one results to inform the development of phase two question topics. Consistency between focus group performance was ensured through the primary researcher acting as the moderator for both focus groups.

The questionnaire lower response rate (28%) may be in part explained by technical issues including a period of down time due to contractual issues with HEE and the host website, thus leading to the need for the survey link to be changed and re-sent to trainees. While every effort was made to mitigate this, there is the possibility that potential hierarchical relationships existed between participants who ranged in seniority, thus influencing the contribution of more junior individuals.

With regard to the transferability of the study's findings, several factors may be considered by the reader who can establish whether the results may be transferable to their own educational context. Purposive sampling was utilized to maximize the homogeneity of participants by using physicianly specialties. Consequently, a lack of parity of assessments and ePortfolio engagement across non-physicianly specialties may impair transferability. Participants were selected due to their experience of both the traditional WPBA and newer SLE. It is possible that individuals who have not experienced both assessment forms may have different perceptions of the value SLEs play in the identification of training difficulties. The study provides data from one geographical region (NW England) and further testing of the direct applicability of these findings to other UK regions and to postgraduate medical training in other countries is therefore warranted.

Conclusion

Our findings provide a novel insight into physician trainees' perceptions of the value of SLEs in the assessment of clinical competency. While the SLE introduction has appeared to enhance the TiD identification process, our findings highlight that further improvement could be made. Perhaps the most significant area for progress includes ensuring sufficient investment from both the trainee and assessor, in order to maximize the potential benefits of the assessment process. Through standardization of assessor training, revalidation and encouragement of a more active role in the selection and management of topics, as well as trainee empowerment to engage with formative elements, optimal SLE performance may be achieved. Findings may be of interest to those involved in postgraduate medical training, including Training Programme Directors and Educational Supervisors, as well as those involved in the remediation process.

Future research may wish to consider a more longitudinal follow-up of trainees' perceptions through their training to assess if engagement with the assessment process evolves as training progresses. Additional studies involving participants across training regions may enhance the transferability of findings. It may be valuable to conduct research into other medical specialties such as surgery and primary care to establish if these findings translate to trainees from different specialty backgrounds.

References

[1] Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. Br Med J 2010; 341:c5064.

[2] Driessen E, Scheele F. What is wrong with assessment in postgraduate training? Lessons from clinical practice and educational research. Med Teach 2013; 35(7):569–74.

[3] Postgraduate Medical Education and Training Board. Workplace based assessment: a guide to implementation, 2009. Available via https://faculty.londondeanery.ac.uk/other-resources/files/PMETB_WPBA_Guide.pdf (Accessed 14 July 2018).

[4] Van Der Vleuten CPM, Verhoeven B. In-training assessment developments in postgraduate education in Europe. ANZ J Surg 2013; 83:454–9.

[5] Holmboe ES. Work-based assessment and co-production in postgraduate medical training. GMS J Med Educ 2017; 34(5):58.

[6] Royal College of Physicians and Surgeons of Canada. CanMEDS. Royal College of Physicians and Surgeons of Canada, Ottawa, Canada, 2015. Available via http://canmeds.royalcollege.ca/uploads/en/framework/CanMEDS%202015%20Framework_EN_Reduced.pdf (Accessed 9 September 2018).

[7] Tan J, Tengah C, Heng Chong V, Liew A, Naing L. Workplace based assessment in an Asian context: trainees' and trainers' perception of validity, reliability, feasibility, acceptability, and educational impact. J Biomed Educ 2015; 1–8.

[8] Collins J. Foundation for excellence; An evaluation of the foundation programme. Medical Education England. Oxford, UK, 2010. Available via http://cmec.info/wp-content/uploads/2014/07/Foundation-for-Excellence-An-evaluation-of-The-Foundation-Programme-The-Collins-Report.pdf (Accessed 8 August 2018).

[9] Yeates P, O'Neill P, Mann K. Seeing the same thing differently: mechanisms that contribute to assessor differences in directly-observed performance assessments. Adv Health Sci Educ 2013; 18(3):25–41.

[10] Kessel K, Jenkins J, Neville E. Workplace based assessments are no more. Br Med J 2012; 345:e6193.

[11] Brown TRP, Beggs KL, Haig A. Trainee doctors' switch to Supervised Learning Events: The effect on form submission timing. 17th Annual International Association of Medical Science Educators (IAMSE), St Andrews, UK, 2013.

[12] Joint Royal Colleges of Physicians. Recommendations for specialty trainee assessment and review incorporating lessons learnt from the workplace-based assessment pilot, 2014. Available via https://www.jrcptb.org.uk/sites/default/files/April%202014%20Recommendations%20for%20specialty%20trainee%20assessment%20and%20review.pdf (Accessed 8 September 2018).

[13] Paice E, Cox J, King J, Hutchinson A. Understanding doctors' performance. Radcliffe Publishing, Oxford, UK, 2006.

[14] Patel M, Agius S, Wilkinson J, Patel L, Baker P. Value of supervised learning events in predicting doctors in difficulty. Med Educ 2016; 50(7):746–56.

[15] Yao DC, Wright SM. The challenge of problem residents. J Intern Gen Med 2001; 16:486–92.

[16] Steinert, Y. The problem junior: whose problem is it? Br Med J 2008; 336:150–3.

[17] National Association of Clinical Tutors. Managing trainees in difficulty (version 3). Practical advice for educational and clinical supervisors. NACT, UK, 2013. Available via http://www.gmc-uk.org/Final_Appendix_5_Trainees_in_Difficulty.pdf_53816759.pdf (Accessed 8 July 2018).

[18] Steinert Y. The "problem" learner: whose problem is it? AMEE Guide No. 76. Med Teach 2013; 35(4):e1035–45.

[19] Mitchell C, Bhat S, Herbert A, Baker P. Workplace-based assessments in Foundation Programme training: do trainees in difficulty use them differently? Med Educ 2013; 47(3):292–300.

[20] Naidoo S, Lopes S, Patterson F, Mead HM, Macleod S. Can colleagues', patients' and supervisors' assessments predict successful completion of postgraduate medical training? Med Educ 2017; 51:423–31.

[21] Glaser BG, Strauss A. The discovery of grounded theory.Aldine, Chicago, IL, 1967.

[22] Willig C. Introducing qualitative research in psychology. 3rd edition, McGraw-Hill, London, UK, pp 69–81, 2013.

[23] Online Surveys, 2017. Available via https://admin.onlinesurveys.ac.uk/accounts/register/organisation (Accessed 8 September 2018).

[24] Great Britain. Data Protection Act. Department for Digital, Culture, Media & Sport, London, UK, 2018.

[25] Glaser BG. Theoretical sensitivity. The Sociology Press, Mill Valley, CA, 1978.

[26] Rees CE, Cleland JA, Dennis A, Kelly N, Mattick K, Monrouxe LVB. Supervised learning events in the Foundation Programme: a UK-wide narrative interview study. Br Med J Open 2014; 4:10.

[27] Basu I, Parvizi S, Chin K. The perception of online work-based assessments. Clin Teach 2013; 10(2):73–7.

[28] Massie J, Ali JM. Workplace-based assessment: a review of user perceptions and strategies to address the identified shortcomings. Adv Health Sci Educ 2016; 21(2):455–73.

[29] Hunter AR, Baird EJ, Reed MR. Procedure-based assessments in trauma and orthopaedic training—the trainees' perspective. Med Teach 2015; 37(5):444–9.

[30] Crossley J, Jolly B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. Med Educ 2012; 46(1):28–37.

[31] Sowden D, Hinshaw K. The trainee in difficulty: a viewpoint from the UK. Obstet Gynaecol 2011; 13:239–46.

[32] Saedon H, Salleh S, Balakrishnan A, Imray CHE, Saedon M. The role of feedback in improving the effectiveness of workplace based assessments: a systematic review. BMC Med Educ 2012; 12:25.

[33] Gaunt A, Patel A, Rusius V, Royle TJ, Markham DH, Pawlikowska T. 'Playing the game': how do surgical trainees seek feedback using workplace-based assessment? Med Educ 2017; 51(9):953–62.

[34] Mukhtar O, Griffin J, Naheed S, Bryant C. The multiple consultant report: new assessment; same old problems. Clin Teach 2018; 15(1):34–7.

[35] Brown JM, Lowe K, Fillingham J, Murphy PN, Bamforth M, Shaw NJ. An investigation into the use of multi-source feedback (MSF) as a work-based assessment tool. Med Teach 2014; 36(11):997–1004.

[36] Johnson S, Cai A, Riley P, Millar LM, Mcconkey H, Bannister C. A survey of Core Medical Trainees' opinions on the ePortfolio record of educational activities: beneficial and cost-effective? J R Coll Physicians Edinb 2012; 42:15–20.

Appendices

Appendix A: Online Questionnaire

Online Questionnaire: Trainees' perceptions of workplace-based assessments in assessing clinical competence and do they identify trainees in difficulty?

Performance to date

1. Have you had an ARCP outcome other than outcome 1 to date?

    1.1. If so, at what stage of training and what were the reasons?

    1.2. Had feedback on your WPBAs reflected this outcome?

    1.3. Did the feedback on your Educational Supervisor's report reflect this outcome?

Supervised Learning Events

2. Please rate the following Supervised Learning Events in order of usefulness in identifying clinical competence add in ACATs as well

3. Following the changes made to WPBAs, do you feel the newer SLE forms more accurately record clinical competence? If yes, why; if no, why not?

4. How do you think WPBAs could be further improved to help assess clinical competence?

5. Following the changes made to WPBAs, do you feel the newer SLE forms help with identifying trainees in difficulty earlier? If yes, why; if no, why not?

6. How do you think WPBAs could be further improved to help identify trainees in difficulty?

Experience of completing WBPAs for other trainees

7. Think of the assessments you have completed for other colleagues over the last year. Did you feel the completion of the assessment accurately recorded competence of the trainee?
   If yes, please expand. If not, how could the process be improved?

Professionalism and performance

8. How useful do you think WPBAs are at assessing professionalism? If useful, how? If not, how could they be improved?

	Case Based Discussion	Mini-cex	Multi source feedback	Directly Observed procedure	Multi consultant report	Educational supervisor report	ACAT
Very Useful							
Somewhat Useful							
Not very useful							
Not at all useful							

Appendix B: Focus Group Schedule

Focus Group Schedule

Preamble: Thank you for kindly agreeing to talk to us about your experience of Supervised Learning Events and their value in assessing clinical competency and identifying trainees in difficulty. We will not be recording or discussing specific clinicians or clinical trusts. In 2014, SLEs were introduced to the specialty trainee e-portfolios.

1. **Firstly, we will discuss your opinions of SLEs in assessing clinical competency.**

Prompt questions:

- In your opinion, how do they differ from traditional WBA forms pre 2014?
- Which aspects of the assessments are most beneficial in assessing clinical competency?
- Which aspects of the assessments are least beneficial in assessing clinical competency?
- Which assessments are of most benefit in establishing clinical competency?
- To what extent does timing of completion of SLE forms influence their value?

- To what extent does the assessor influence SLEs value of assessing clinical competency?
- How could SLEs be improved in order to assess clinical competency?

2. **Next, we will discuss your opinions on SLEs value in identifying trainees in difficulty.**

   Prompt questions:

   - How are SLEs used to help identify trainees in difficulty?
   - Which SLEs in your opinion are most valuable in establishing trainees in difficulty?
   - To what extent does the assessor influence SLEs value in identifying trainees in difficulty?
   - How do you think SLEs could be improved to help identify trainees in difficulty earlier?

3. **Finally, we would like to discuss your experiences of completing assessments for other colleagues. Consider the assessments you have completed for others in the last year.**

Prompt questions:

- To what extent did you feel you could be honest when completing the assessment for trainees?
- How did your feedback reflect your opinions of the trainees' clinical competency?
- What (if any) are your experiences of trainees "masking" areas of weakness in the completion of assessments?
- Have you ever considered a trainee to select you as an assessor in circumstances when you were not the most appropriate assessor?
- How (if any) could completion of assessments for other colleagues be improved to assess clinical competency?

Finally, are there any other aspects of SLEs value in assessing clinical competency and identification of trainees in difficulty that you wish to discuss further?

Thank you for participating in this focus group.