



Minerva Access is the Institutional Repository of The University of Melbourne

**Author/s:**

Scarff, CE;Corderoy, RM;Bearman, M

**Title:**

In-training assessments: 'The difficulty is trying to balance reality and really tell the truth'

**Date:**

2018-02-01

**Citation:**

Scarff, C. E., Corderoy, R. M. & Bearman, M. (2018). In-training assessments: 'The difficulty is trying to balance reality and really tell the truth'. *Australasian Journal of Dermatology*, 59 (1), pp.e15-e22. <https://doi.org/10.1111/ajd.12555>.

**Persistent Link:**

<https://hdl.handle.net/11343/292241>

Title page

Title:

In training assessments: “The difficulty is trying to balance reality and really tell the truth”.

Short running title:

In training assessments.

Authors:

Catherine E Scarff (corresponding author)

MBBS, MMed, MHPE, FACD

HealthPEER – Health Professions Education and Educational Research

Monash University

Building 13C G08

Clayton Vic 3168

Australia

Ph: +61 3 9905 0090

cscarff@mac.com

Robert M Corderoy

JP, BA, MSc, MEd, PhD, FACE

Director: Educational Development, Planning and Innovation

Australasian College of Dermatologists

Rhodes, NSW, Australia

Margaret Bearman

PhD

Associate Professor

HealthPEER – Health Professions Education and Educational Research

Monash University

Clayton, Victoria, Australia

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/ajd.12555](https://doi.org/10.1111/ajd.12555)

This article is protected by copyright. All rights reserved

Received Date : 04-Jun-2016

Revised Date : 31-Jul-2016

Accepted Date : 03-Aug-2016

Article type : Original Research

## **Title**

In training assessments: “The difficulty is trying to balance reality and really tell the truth”.

## **Abstract**

### Background

In training assessments (ITAs) are well placed to evaluate trainees' progress and give valuable feedback on their performance. Many factors have the potential to impact on supervisors during the process of their completion of assessments and these can affect the final results recorded.

### Methods

This is the second part of a study of supervisors of the Australasian College of Dermatologists (ACD) and presents the qualitative data captured on their opinions regarding the ACD in training assessment process and influences upon their ITA ratings.

### Results

Supervisors highlighted the benefits of this assessment tool, and many limitations were also noted. Potential influences upon supervisor ratings included the relationship between the supervisor and trainee and the level of honesty in completion and delivery of the assessment.

### Conclusions

Many factors influence supervisors in the completion of the ITAs. These include the impact of interpersonal relationships and concerns about the consequences of delivering a negative assessment which sometimes lead

supervisors to modify the assessment they deliver to the trainee. Further research is needed into honesty in assessment judgements.

## **Background**

Assessment is an important part of any educational system. In addition to documenting the achievement of standards, assessment can provide valuable feedback to trainees and guide their future learning (1). Performance assessments, such as in training assessments (ITAs) fulfil both these purposes. ITAs cover a broad range of domains of performance, and have the ability to assess professional attributes along with clinical knowledge and skills (for example (2)). As they are based on a period of observation of trainees in their real-world work setting, they are ideally suited to provide detailed feedback. In summary, they provide evidence of what the trainee actually “does” with their knowledge and skills in practice, in contrast with what trainees “know”, as in a written exam, or can “show”, for example in an Objective Structured Clinical Examination (OSCE) setting (3).

While ITAs are a valuable assessment tool, as with any performance-based assessment there are many factors that may potentially compromise or confound their validity (4). Some relate to the design and use of the tool itself, including logistical issues (5), the scale design (6) and training in use of the tool (1,7). Other general performance-based assessment concerns include the number of samples of performance considered (4), assessor biases (8), level of expertise in evaluation (9) and the mental workload involved in assessment (10).

The clinical supervisors who complete ITAs are in a potentially difficult position. They often occupy multiple conflicting roles including mentor, teacher and supporter as well as assessor and gate-keeper (11–13). Supervisors often have limited time to make their assessments. This, together with concerns about trainees’ reactions to an assessment and limitations in the training they have received, may contribute to problems with the assessment process (14).

The Australian College of Dermatologists (ACD's) assessment of trainees includes a summative ITA (SITA); the current format has been in use since 2006. The process of SITA completion in the ACD are described in detail elsewhere (15). In summary, the SITA is a competency-based tool which is performed six-monthly on all trainees. The current tool has a combination of a numerical scale, where domains of performance are rated on a level of competency scale 1 to 5, and a descriptive scale, where supervisors record comments regarding the trainee's strengths and areas where they require development. In this setting, they are completed by three assessors at a single point in time, though in others, ITA may be completed by only one assessor (for example (12,16)).

This article presents the second part of a broader survey study examining the use of the SITA in the Australasian College of Dermatologists (ACD). This study is the first to document and describe supervisors' opinions of the SITA process with a view to highlighting issues for future intervention as required. The quantitative results have been presented and (15) indicated that, while supervisors agreed with the need to provide trainees with feedback and report underperformance when present, they noted limitations in their ability to do so with the SITA tool. Supervisors did not have complete confidence in their judgements, as noted by the number who wished to change a previously submitted assessment. Responses also indicated concerns about the impacts of negative assessments, for example 45% of respondents agreed that 'a negative comment from me may come back to haunt me.' However, these results only quantify concerns, they do not provide insight into the thinking behind the responses. Qualitative data provide a richness and exploration of the "why" and "how" of "real life" behaviour" compared with the "how much" which is presented in the previous paper p.4 (17).

In this part of the study, we sought to qualitatively describe the perspectives of supervisors in regards to the factors that influence their assessments of Australian dermatology trainees. Our research questions were:  
"What are ACD supervisors' opinions of the SITA?"

What are ACD supervisors' perceptions of influences upon their ratings?"

## **Methods**

### *Survey design*

No pre-existing surveys were found which related to the proposed topic, so a survey was designed by the authors to explore supervisors' perspectives regarding SITA. Development details, including piloting, are described in Scarff, Bearman and Corderoy (15). In summary, the questionnaire was developed on key themes, primarily derived from a comprehensive literature review (the mapping of survey questions to the literature is available on request from the authors). Design of the survey items and distribution of the survey were based on methods described in Mertens (18). Questionnaire development occurred over several iterations between all authors and included input from additional staff of the Faculty of Medicine, Nursing and Health Sciences, Monash University. The survey was anonymous, and presented in an online platform (Qualtrics survey software) to facilitate distribution, completion and collection of data. The survey was open to all Fellows of the ACD who self-identified as either a clinical supervisor (CS) or supervisor of training (SOT).

The survey contained 67 items, with the 17 free-text prompts relevant to this paper reported in Table 1.

### *Qualitative analysis*

All free-text answers were extracted from the survey for analysis. Two authors (CS and MB) performed the qualitative data analysis using a thematic analysis technique (19). This involved each reading the responses to the free-text questions in the survey several times. Each then independently 'coded' this data, that is labelled text using words or short phrases. One author (CS) clustered these codes into themes and then clustered these themes against the research questions. Themes were discussed (CS and MB) until consensus was achieved. A small number of codes, which were not relevant to either research question, are not presented. These included themes which relate to supervisors' ideas for modifications to the SITA or opinions on why the SITA is conducted.

*Ethical approval*

The study design was approved by Monash University Human Research Ethics Committee (CF 13/1450 – 2013000767 27 May 2013).

<b>“What are ACD supervisors’ opinions of the SITA and what are their perceptions of influences upon their ratings?”</b>	<b>Number of responses to question</b>
<b><u>Logistics</u></b>	
Have there been logistical issues that have affected your completion of a SITA? If so, please describe.	15
<b><u>Usefulness of the tool</u></b>	
The SITA effectively captures the performance of the Trainee. Please elaborate on the reasons for your response.	28
The SITA is essentially a formality. Please elaborate on the reasons for your response.	29
In what ways do you think the SITA is most useful for the Trainee?	33
In what ways do you think the SITA is least useful for the Trainee?	30
<b><u>Confidence in my own judgement</u></b>	
If there have been times when on reflection, you would have liked to change a score you gave a Trainee, please describe the reasons leading to your wish to change.	19
Please describe how the SITA supports your judgment in identifying underperforming trainees.	23
Please describe how the SITA supports your judgment in identifying excellent trainees.	25

Please describe the SITA supports your judgment in identifying borderline trainees.	26
<b><u>Relative use of the SITA</u></b>	
The reasons I rarely score a Trainee a 5 are:	14
The reasons I rarely score a Trainee 1 or 2 are:	27
<b><u>Interpersonal relationship to the trainee</u></b>	
If you can think of one, please describe a situation where you feel the SITA assessment you gave was influenced by your negative relationship with the Trainee.	16
If you can think of one, please describe a situation where you feel the SITA assessment you gave was influenced by your positive relationship with the Trainee.	15
If a trainee has tried to influence your assessment of them previously, please describe the situation.	4
If you agree that making a negative comment may adversely affect your relationship with the trainee, please comment on how, for example create tension in clinic, risk of litigation etc.	18
If you agree that your assessment would differ if you had to deliver it to the trainee directly, please describe the ways in which your assessment would differ.	5
If a trainee has reacted negatively to your assessment in the past, please describe how a Trainee.	8

**Table 1: Qualitative questions asked of supervisors in the survey**

## Results

The completion rate for the survey was 41% (n=32/79 respondents).

However, 72/79 (91%) of all CS and SOT opened the survey and submitted

responses to some of the total of 67 questions. The gender and geographic distribution of respondents who completed the survey is listed in table 2.

(n)	NSW	Vic	SA	WA	Qld	Total
<b>M</b>	19% (6)	16% (5)	9% (3)	3% (1)	6% (2)	53% (17)
<b>F</b>	9% (3)	25% (8)	3% (1)	3% (1)	6% (2)	47% (15)
<b>Total</b>	28% (9)	41% (13)	13% (4)	6% (2)	13% (4)	100% (32)

**Table 2: Distribution of survey respondents gender and state of practice. There were no respondents from Tasmania, the ACT or the Northern Territory.**

Not all respondents gave qualitative explanations to the quantitative data. For example, while 23 respondents agreed that a negative comment from them could affect their relationship with the trainee, only 18 gave an explanation for this in the free-text area. The thematic analysis indicated a broad range of themes, as described in Table 3. Each theme will be described, illustrated with relevant quotes from the survey data.

Research question: What are ACD supervisors' opinions of the SITA?
<p><b>Category: Strengths of the SITA process</b></p> <p>Themes: Acknowledgement of performance  Documentation of progress  Broad coverage of tool  Facilitation of planning for the future</p> <p><b>Category: Limitations of the SITA process</b></p> <p>Themes: Not helpful for excellent performers  Design constraints  Challenges in making judgements  Limited or no effect on trainee performance</p>
Research question. What are ACD supervisors' perceptions of influences

upon their ratings?
<p>Themes: Misunderstandings of the process</p> <ul style="list-style-type: none"> <li>Factors relating to the ACD</li> <li>Trainee's personality and/or behaviours</li> <li>Interpersonal relationship between supervisor and trainee</li> <li>Consequences of a negative SITA</li> <li>Methods for minimising issues with delivering a negative SITA</li> <li>Level of honesty in SITA completion</li> </ul>

**Table 3: Themes reported by supervisors to the free-text questions in the survey**

### **Themes regarding supervisors' opinions of the SITA**

These are reported under two categories: a category of themes that describe the strengths of the SITA and a category of themes that describe limitations of the SITA.

#### *Strengths of the SITA process*

Respondents reported that the SITA provides the ability to recognise and acknowledge good performance, as well as highlight those not performing well. The regular, compulsory and formal feedback opportunity provided by the SITA was noted; promotion of a dialogue about performance was included within this theme.

The documentation of trainees' progress throughout the program was seen as important and the framework of the SITA was noted by some to facilitate the assessment of a wide range of domains, however as will be discussed in the next section, others expressed a differing view.

The strength and importance of the tool in focusing development and planning for the future for poorly performing learners was noted. See table 4.

<u>Themes</u>	<i>Illustrative quote</i>
<u>Acknowledgement of performance</u>	<p><i>"to encourage and recognise good performance"</i></p> <p><i>"an opportunity to highlight underperforming registrars".</i></p> <p><i>"I think it's a fantastic tool for giving feedback to trainees"</i></p>
<u>Document trainee's</u>	<i>"Provides important evidence of a trainees success"</i>

<u>progress</u>	<i>or failure in training”</i>
<u>Broad coverage of tool</u>	<i>“covers a broad range of skills we want registrars to have”.</i>
<u>Facilitation of planning for the future</u>	<i>“Gives focus on areas to improve”</i>

**Table 4: Themes relating to the strengths of the SITA**

#### *Limitations of the SITA process*

Some respondents felt the tool was not as beneficial for trainees who are performing well compared with those performing poorly.

Issues relating to both the design of the tool and of the overall SITA process were noted. Some respondents highlighted issues of time in relation to SITA completion, including noting the time the assessment takes to complete and conflicts with other pressures on supervisors’ time. Comments were also made about how the SITA constrained feedback and that the tool did not necessarily encapsulate actual work practices.

The limitations of the ratings component and the need for supervisors to document comments in particular was reported. Respondents wrote of the challenges in making their judgement themselves, including their uncertainty in determining the level at which to assess trainees and in relation to the potential limitations of a single assessor judgement.

Some identified perceived barriers to making effective judgements and the need to discuss trainees with other supervisors. There were also some concerns about the level of confidentiality of the assessments given.

The ability of the SITA to have an effect on trainee performance and outcomes was also questioned with some stating it provides no learning goals. It was sometimes described as a ‘mandated’ rubber-stamping exercise.

Others noted the need for trainees to be receptive to the feedback that they are given and some felt that trainees were already well aware of their level of performance. See table 5.

Themes	Illustrative quotes
--------	---------------------

<p><u>Not helpful for excellent performers</u></p>	<p><i>“an excellent trainee does not need a SITA assessment, really”.</i></p>
<p><u>Design constraints</u></p>	<p><i>“We are offered limited choices...”</i></p> <p><i>“Many of the parameters measured in the SITA are not captured with the way clinics are run”.</i></p> <p><i>“The only reason the SITA...captures information about the candidate is because if [sic] the comments section”.</i></p>
<p><u>Challenges in making judgements</u></p>	<p><i>“difficult as an assessor to determine if trainee is being assessed at the right level”</i></p> <p><i>“the only way to really know is to observe and work with the trainee and discuss the trainee with other colleagues”</i></p> <p><i>“I am unconvinced that confidentiality is all that good currently”.</i></p>

<p><u>Limited or no effect on trainee performance</u></p>	<p><i>“It is often a tick box formality by most CSs and some SoTs. It does not set clear learning goals.”</i></p> <p><i>“It results in certain areas being "ticked off" in an organisational and PC sense that require more subtle and discriminating examination. I feel it is both a poorly focussed tool and a missed opportunity.”</i></p> <p><i>“We only do it because the current AMC mandates it for transparency etc”.</i></p> <p><i>“Trainees need to be able to take negative feedback if they are to improve after assessment.”</i></p> <p><i>“it serves little purpose. First most are aware if there [sic] behind the 8th ball”</i></p>
---	--

**Table 5: Themes relating to the limitations of the SITA**

### **Themes regarding supervisors' perceptions of influences upon their ratings**

Some comments indicated misunderstandings of the SITA process are present such as the ratings that will precipitate a negative SITA. Other comments revealed misunderstandings in regards to who was responsible for distribution and collection of forms.

Factors relating to the ACD which impact assessments were noted by some, such as the size of College and long-term relationships between trainees and assessors.

The trainee's personality and behaviours were seen as potential influences on the assessment process, for example the influence of affability was noted by some and a direct attempt to influence an assessment in advance in one case.

The interpersonal relationship between the trainee and assessor was also noted as a factor influencing assessment, with both a positive and negative relationship having an impact at times. However, this was not universal, with

other respondents noting that neither trainee personality nor their interpersonal relationships impacts on their assessment.

Some respondents noted that the consequences of a negative assessment could effect their assessments. Some referred to the potential impact on trainees and a reluctance to upset or cause additional stress. Caution in the way feedback is delivered and the need to be sensitive was mentioned by some.

Other comments related to the potential of a negative assessment to impact on an assessor, for example, in terms of the working environment or by reflecting on their own ability as a supervisor. Some spoke of their own previous experiences with giving a negative assessment.

While some respondents indicated that they have no particular concerns delivering their feedback others described methods they have used for minimising issues with delivering a negative SITA. This included giving feedback verbally or writing a comment instead of assigning a lower score.

Some respondents suggested that some supervisors do not deliver their assessment judgement with complete 'honesty' and it was noted by some that there may be consequences to this lack of honesty including trainees with uncorrected deficits presenting for the Fellowship exam.

However others felt the anonymity with the SITA process permitted more 'honesty', as indicated by comments from some supervisors in relation to how their assessments would change if they had to deliver them directly to the trainee. See table 6.

<u>Themes</u>	<i>Illustrative quotes</i>
<u>Misunderstandings of the process</u>	<i>"If I give a score lower than 4 in all, it will be seen as unsatisfactory, that will affect Trainees [sic] future"</i>
<u>Factors relating to the ACD</u>	<i>"I do not wish to impede the progress I [sic] registrars especially knowing that I will be working alongside them one day as fellow dermatologists."  "It being a small college , CSs do not want to 'make enemies'."</i>

<p><u>Trainee's personality and/or behaviours</u></p>	<p><i>"A registrar who is punctual, enthusiastic, reliable, personable and helpful will be more likely to score highly than those that seem less interested."</i></p> <p><i>"a trainee stated that they didn't want to get any borderlines as they were shocked to have received them at the previous assessment"</i></p>
<p><u>Interpersonal relationship between supervisor and trainee</u></p>	<p><i>"I have given the trainee [an] outstanding 5 score, but possibly my relationship with them influenced that."</i></p> <p><i>"I personally liked the trainee, but (they were) definitely below standard in performance, however I upgraded (their) SITA"</i></p> <p><i>"Even though (the trainee) wasn't doing anything unsafe, my opinion of (them) did influence some of my points on (the) SITA"</i></p> <p><i>"I had a registrar who I personally did not like (just didnt [sic] gel I guess), but they were above average in their academic and professional standard, and I marked them accordingly."</i></p>

Consequences of a  
negative SITA

Author Manuscript

*“put extra stress on trainee”*

*“[the] difficulty is trying to balance reality and really tell the truth with underperforming trainees & to not to be 'too harsh and mean' and trying to be encouraging and supportive of [trainees]”*

*“A negative SITA certainly is not good for the working climate”*

*“I can cope with a trainee not liking me because I marked them down but threat of litigation worries me.”*

*“For most trainees it is difficult to give unsatisfactory performance, in some cases doing so could also reflect back on a failing of the CSs to help the trainee obtain an adequate SITA”*

*“I am afraid of upsetting the Trainee. I have been in this situation in the past and it was not pleasant”*

*“Blatantly rude...many aggressive personal comments. Text messages late at night...subtle comments, avoidance of work behaviour etc. I ALWAYS think twice about a negative SITA”*

*“Avoiding me in clinic Not asking [for] advice on patients and management from me”.*

<p><u>Methods for minimising issues with delivering a negative SITA</u></p>	<p><i>“While a negative report that a trainee becomes aware of may have a negative effect on the professional relationship from their point of view, from mine I believe I am mature and experienced enough to rise above that and can maintain the trainer/trainee relationship without significant effect.”</i></p> <p><i>“I have at times informally spoken to candidates about their issues whilst still giving them "borderline" rather than "unsatisfactory"”</i></p> <p><i>“If i give the trainee the benefit of a 'wavering' score by scoring a 3 instead of a 2 then i make sure the area that needs improving is clearly listed in the comments section. This is sometimes a useful strategy in dealing with trainees with delicate/ fragile egos that is borderline.”</i></p>
<p><u>Level of honesty in SITA completion</u></p>	<p><i>“Most Css fail to give an 'honest' SITA .They may complain about a trainee, but the SITA form is usually satisfactory”</i></p> <p><i>“the reluctance of supervisors to award negative comments results in many trainees [attempting] the exams with knowledge and behavioural defects”</i></p> <p><i>“more likely to sugar coat any criticisms”</i></p>

**Table 6: Themes regarding supervisors' perceptions of influences upon their ratings**

## **Discussion**

This study is the first to qualitatively document ACD supervisors' opinions on the SITA process. It describes strengths, limitations and perceived influences on the process. Many supervisors feel the strengths of the SITA are in providing a valuable opportunity to discuss performance with the trainee across a broad range of activities. They note the SITA's potential to provide direction and identifying underperforming and exceptional trainees. Supervisors' opinions of limitations included: the design of the tool; its ability to both reflect the way training occurs and how clinics run; and how well poorly performing learners could be identified. Influences on supervisors ratings were included: some misunderstandings about processes; the ACD as a context; the trainee; the interpersonal relationship between supervisor and trainee; consequences to giving a negative SITA; managing the negative SITA message; and perceptions of honesty in SITA completion.

The qualitative data presented here provide explanation to the quantitative data obtained. For example, we previously published that over 60% of responding supervisors had reflected a desire to change a previously submitted SITA score. Here, the reasons for this are revealed and include lack of input from others who work with the trainee, fear of upsetting the trainee and disagreement with the possible consequences of a given assessment result. Our results show that for some supervisors, there are significant issues with the process, which may currently be impacting on the results recorded. It is clear that some supervisors are currently not giving trainees their true judgement as seen by the methods for minimising issues with a negative SITA reported by some supervisors. Moreover, these results indicate that some supervisors have a level of dissatisfaction and even hostility with the process. This lack of engagement with the process may impact on the quality of the assessments submitted, if their completion is regarded as merely a "formality" (1).

## **Misunderstandings and the impact of College**

Responses from some supervisors showed misunderstandings about the process of form completion and the implications of certain ratings, specifically those that potentially lead to an unsatisfactory SITA. These issues are amenable to the education and training of supervisors, however, those that indicated that the small size of the ACD and ongoing relationships between those involved impact their assessment decisions are more challenging to address. This is a particularly noteworthy finding as this study's setting of a smaller specialist assessment community has not previously been explored in the literature, and highlights some of the challenges faced by work-based assessments in these particular environments.

## **Relationship with the trainee**

Trainee personality and behaviours clearly affect the SITA process for some supervisors. While not universal, some supervisors agreed that it can be difficult to prevent their personal feelings about a trainee from influencing their assessment and gave examples of when this had occurred. Studies from organisational psychology support this phenomenon (20), which may be ameliorated with consideration of the rating context, again including rater training (11). However, other options such as attention to who completes and the delivers assessments to the trainee are worthy of consideration.

## **Fear of being negative**

As presented, our respondents reported several explanations for being fearful of giving negative assessment results, and many of these have been reported in the literature. Some were concerned that they would cause stress for the trainee and other authors agree, noting a negative comment may be discouraging and result in "demotivating students, who will subsequently perform at substandard levels" (p.248) (11). Some respondents wrote of their first hand experiences of giving a low ratings, while for others, fears and concerns about the consequences of giving their private judgement for themselves were described. This included the potential for a negative

assessment to reflect back on the supervisor and this too has been reported in other health professional education settings (21,22).

The behaviour of trainees is seen as a potential influence on the supervisor, with a small number of supervisors reporting significant negative responses from trainees arising from a negative assessment of them. The fact that some report that this may influence their future handling of assessment decisions is of significant concern. This together with the practise of trainees attempting to directly pressure supervisors have also been reported previously (23,24) and brings into question the validity of the assessment. Robust procedures to deal with trainees who attempt to influence or intimidate supervisors need to be explored.

### **Methods of dealing with difficult assessment decisions**

Some respondents described the conflict between what they wanted to say or write and the potential outcome of doing so and other authors have reported that a supervisor's public assessment result may not always be the same as their private one (25). Several responses implied or stated that assessors may avoid delivering their true judgement if the cost of doing so will be too great either for the trainee or the assessor. Again, these costs were seen to be both immediate in terms of the anticipated reaction of the trainee and long-term, where the small size of the ACD and the need for ongoing relationships with Fellows may be relevant. Some respondents described methods of manipulating the assessment process in order to avoid potentially difficult situations by giving a higher rating than the one they feel is justified, but providing their true judgement in a written or verbal comment to the trainee (table 6). This avoids negative consequences arising from the assessment, but the supervisor feels that they have communicated their judgement, advice and feedback. Again, the lack of correlation between some comments and ratings and across different feedback mediums has also been reported in the literature (13,26).

Future directions

Supervisors require adequate training and support in order to conduct assessments. Training needs to include not only the practical aspects of the process of assessment and giving feedback, but also the importance of and obligation to provide an honest appraisal of the trainee's performance. Support for supervisors must be clear, being mindful that these assessment duties are added onto the workload of already busy clinicians. Allowing supervisors to discuss ratings with others who work with the trainee was mentioned by respondents and if presented as a group consensus, may avoid many of the concerns reported in this study. This practice is used in several other settings, for example Dermatology training in the UK (27).

In relation to significant reactions of trainees to assessments that some respondents reported, the support of both their colleagues and the ACD is important, but clear policy to deal with this situation is also required. In our earlier article, we discussed some options; a robust procedure to deal with trainees who attempt to influence or intimidate supervisors needs to be explored.

By their ability to assess trainees in their real-world situation, the SITA should be at the top of the assessment tree. However, the issues highlighted in this study are currently limiting their ability to achieve this.

### **Strengths, limitations and future directions**

This study has highlighted many areas of complexity in the performing and delivering of performance assessments to trainees. It has done so through a national study of a specialist college, which afforded respondents the opportunity for anonymous and hence more likely to be honest responses. This data highlights the human dimension of this complex process (28), which has received less attention to date (21).

Limitations to this study are several. Online surveys are reported to have lower response rates than paper-based ones (29). Reminders to participate in the study and incentives were employed to maximise response rate and the rate of 41% is seen as acceptable (29). Other limitations include the format

which, while allowing anonymous response, constrains the length of description.

Most notably, response to surveys such as these, tend to come from those with strong views, which may also be more negative. As a qualitative study, this work does not seek to generalise its findings. Rather, qualitative data describes the experiences of those involved and raises awareness of issues and may point to the need for further research or investigation in certain areas in more depth. Despite small numbers of respondents, this study has certainly identified issues with the current working of the SITA process. Further research is required to examine in greater depth issues relating to the delivery of assessment judgements from the supervisor to the trainee in the performance assessment setting.

## **Conclusion**

Supervisors in this study recognised the benefits of the ITA process but are also alert to its limitations. Several factors which have the potential to impact on supervisor's assessments were highlighted and misunderstandings of the SITA procedure were revealed. Some supervisors' concerns and fears regarding giving negative assessments were described; attention needs to be given to the impact that interpersonal relationships may have on the SITA process.

## **References**

1. Bingham CM, Crampton R. A review of prevocational medical trainee assessment in New South Wales. *Med J Aust.* 2011; 195: 410–2.
2. Ringsted C, Pallisgaard J, Østergaard D et al. The effect of in-training assessment on clinical confidence in postgraduate education. *Med Educ.* 2004; 38: 1261–9.
3. Miller G. The Assessment of Clinical Skills/Competance/Performance. *Acad Med.* 1990; 65: S63–7.
4. Downing SM, Haladyna TM. Validity threats: overcoming interference with proposed interpretations of assessment data. *Med Educ.* 2004; 38: 327–33.

5. Ringsted C, Østergaard D, van der Vleuten CPM. Implementation of a formal in-training assessment programme in anaesthesiology and preliminary results of acceptability. *Acta Anaesthesiol Scand*. 2003; 47: 1196–203.
6. Pangaro L, Ten Cate O. Frameworks for learner assessment in medicine: AMEE Guide No. 78. *Med Teach*. 2013; 35: e1197–210.
7. Dudek NL, Marks MB, Bandiera G et al. Quality in-training evaluation reports—Does feedback drive faculty performance? *Acad Med*. 2013; 88: 1129–34.
8. Williams RG, Klamen DA, McGaghie WC. Cognitive, social and environmental sources of bias in clinical performance ratings. *Teach Learn Med*. 2003; 15: 270–92.
9. Berendonk C, Stalmeijer RE, Schuwirth LWT. Expertise in performance assessment: Assessors' perspectives. *Adv in Health Sci Educ*. 2013; 18: 559–71.
10. Tavares W, Eva KW. Exploring the impact of mental workload on rater-based assessments. *Adv in Health Sci Educ*. 2013; 18: 291–303.
11. Govaerts MJB, van der Vleuten CPM, Schuwirth LWT et al. Broadening perspectives on clinical performance assessment: Rethinking the nature of in-training assessment. *Adv in Health Sci Educ*. 2007; 12: 239–60.
12. Wilkinson TJ, Wade WB. Problems with using a supervisor's report as a form of summative assessment. *Postgrad Med J*. 2007; 83: 504–6.
13. Cohen GS, Blumberg P, Ryan NC et al. Do final grades reflect written qualitative evaluations of student performance? *Teach Learn Med*. 1993; 5: 10–5.
14. Fazio SB, Papp KK, Torre DM et al. Grade inflation in the internal medicine clerkship: a national survey. *Teach Learn Med*. 2013; 25: 71–6.
15. Scarff CE, Bearman M, Corderoy RM. Supervisor perspectives on the summative in-training assessment. *Australas J Dermatol*. 2016; 57: 128–34.
16. Long G. Documentation of in-training assessment for radiology trainees. *Clin Radiol*. 2001; 56: 310–5.

17. Kuper A, Reeves S, Levinson W. An introduction to reading and appraising qualitative research. *BMJ*. 2008; 337: a288.
18. Mertens D. *Research and Evaluation in Education and Psychology: integrating diversity with quantitative, qualitative, and mixed methods*. California: SAGE Publications, 2010.
19. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. *Am J Eval*. 2006; 27: 237–46.
20. Tsui AS, Barry B. Interpersonal affect and rating errors. *Acad Manag J*. 1986; 29: 586–99.
21. Cleland JA, Knight L V, Rees CE et al. Is it me or is it them? Factors that influence the passing of underperforming students. *Med Educ*. 2008; 42: 800–9.
22. Rutkowski K. Failure to fail assessing nursing students' competence during practice placements. *Nurs Stand*. 2007; 22: 35–40.
23. Dudek NL, Marks MB, Regehr G. Failure to fail: The perspectives of clinical supervisors. *Acad Med*. 2005; 80: S84–7.
24. Cacamese SM, Elnicki M, Speer AJ. Grade inflation and the internal medicine subinternship: A national survey of clerkship directors. *Teach Learn Med*. 2007; 19: 343–6.
25. Mohrman AM, Lawler EE. Motivation and performance appraisal behaviour. *Proceedings of the Performance Measurement: Directions for the Future conference*. 1981 Nov 6-8; Dallas, Texas USA.
26. Waung M, Highhouse S. Fear of conflict and empathic buffering: Two explanations for the inflation of performance feedback. *Organ Behav Hum Decis Process*. 1997; 71: 37–54.
27. Joint Royal Colleges of Physicians Training Board. *Specialty Training Curriculum for Dermatology*. London; 2012.
28. Albanese M. Rating educational quality: factors in the erosion of professional standards. *Acad Med*. 1999; 74: 652–8.
29. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? *Assess Eval High Educ*. 2008; 33: 301–14.