2016-05-01

Rising to the challenge: acute stress appraisals and selection centre performance in applicants to postgraduate specialty training in anaesthesia

Roberts, MJ

10.1007/s10459-015-9629-6
ADVANCES IN HEALTH SCIENCES EDUCATION
Springer Science and Business Media LLC

This is a final author's draft of a paper submitted for publication in Advances in Health Sciences Education 2016 DOI: http://dx.doi.org/10.1007/s10459-015-9629-6

Rising to the challenge: Acute stress appraisals and selection centre performance in applicants to postgraduate specialty training in anaesthesia

Martin J. Roberts,[1] Thomas C. E. Gale,[1,2] John S. McGrath,[3,4] Mark R. Wilson[5]

1 Collaboration for the Advancement of Medical Education Research and Assessment (CAMERA), Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK
2 Department of Anaesthesia, Plymouth Hospitals NHS Trust, Plymouth, UK.
3 Royal Devon and Exeter NHS Foundation Trust, Exeter, UK.
4 University of Exeter Medical School, Exeter, UK.
5 College of Life and Environmental Sciences, University of Exeter, Exeter, UK.

Correspondence to:
Martin Roberts
Plymouth University Peninsula Schools of Medicine and Dentistry
Portland Square
Drake Circus
Plymouth, UK
PL4 8AA
(+44)1752 586831
Email: martin.roberts@plymouth.ac.uk

Abstract

Background. The ability to work under pressure is a vital non-technical skill for doctors working in acute medical specialties. Individuals who evaluate potentially stressful situations as challenging rather than threating may perform better under pressure and be more resilient to stress and burnout. Training programme recruitment processes provide an important opportunity to examine applicants' reactions to acute stress. In the context of multi-station selection centres for recruitment to anaesthesia training programmes, we investigated the factors influencing candidates' pre-station challenge/threat evaluations and the extent to which their evaluations predicted subsequent station performance.

Methods. Candidates evaluated the perceived stress of upcoming stations using a measure of challenge/threat evaluation - the cognitive appraisal ratio (CAR) - and consented to release their demographic details and station scores. Using regression analyses we determined which candidate and station factors predicted variation in the CAR and whether, after accounting for these factors, the CAR predicted candidate performance in the station.

Results. The CAR was affected by the nature of the station and candidate gender, but not age, ethnicity, country of training or clinical experience. Candidates perceived stations involving work related tasks as more threatening. After controlling for candidates' demographic and professional profiles, the CAR significantly predicted station performance: 'challenge' evaluations were associated with better performance, though the effect was weak.

Conclusions. Our selection centre model can help recruit prospective anaesthetists who are able to rise to the challenge of performing in stressful situations but results do not support the direct use of challenge/threat data for recruitment decisions.

Key words

Anaesthesia · Challenge and threat · Non-technical skills · Psychological Stress · Specialty training

Introduction

In the acute medical specialties (surgery, anaesthesia, intensive care, emergency medicine, etc.) the ability of doctors to work efficiently and decisively under pressure is of vital importance. Managing acute stress is an essential non-technical skill (Flin et al. 2008) and methods of assessing applicants for selection to training programmes in such specialties need to capture this important aspect of candidate ability. Given the significance of such a high-stakes assessment - success or failure will crucially affect their subsequent medical career – the selection process inevitably places the candidate under considerable pressure. Little is known however, about how junior doctors react to the pressure of undergoing selection processes, or how these reactions might influence performance.

Reactions to pressure may be analysed through the lens of the biopsychosocial model of challenge and threat (Blascovich 2008), a transactional model for understanding individualistic reactions to stress. According to this model, an individual's response to a motivated performance situation is determined by their evaluation of two factors: the demands of the situation and their personal coping resources. Demand evaluations involve the assessment of danger, uncertainty and required effort, whereas resource evaluations involve a self-assessment of the knowledge and skills relevant to the situation (Blascovich and Mendes 2000). Individuals who rate their resources as higher than the demands of the situation (a challenge evaluation) have been found to perform better than individuals who assess their resources to be insufficient to meet the task demands (a threat evaluation) (Seery 2011; Moore et al. 2013; Vine et al. 2015). For example, challenge evaluations prior to a laparoscopic surgery training task predicted both faster completion and more efficient use of visual attention during the task (Vine et al. 2013). In anaesthesia, challenge evaluations have been identified as an important factor in enabling expert anaesthetists to cope with acutely stressful clinical situations and avoid the dangers of chronic stress exposure and burnout in the profession (Larsson et al. 2007; Larsson and Sanner 2010). Research in emergency medicine and general surgery residents has linked challenge/threat appraisals to physiological stress responses during trauma resuscitation simulations (Harvey et al. 2010), while self-reported feelings of stress have been found to correlate negatively with performance during cardiopulmonary resuscitation (Hunziker et al. 2012) and threat evaluations were negatively correlated with clinical reasoning during simulated patient encounters (Pottier et al. 2013).

An individual's evaluation of a potentially stressful situation can be assessed via distinctive patterns of neuroendocrine and cardiovascular response as well as via self-report measures. In naturalistic settings, where real-time evaluation of physiological responses is often impractical, a simple self-report measure of challenge/threat states is the cognitive appraisal ratio (CAR) (Tomaka et al. 1993). A number of studies involving medical trainees have used the CAR (Harvey et al. 2010; Pottier et al. 2011; Pottier et al. 2013; Piquette et al. 2014). The validity of this instrument is supported by known associations between the CAR and expected physiological reactions under stress such as altered heart rate, cardiac contractility, cardiac output and vascular activity (Tomaka et al. 1993; Harvey et al. 2010; Vine et al. 2013). While recent studies have highlighted the importance of challenge/threat evaluations in a variety of medical disciplines, there is a lack of research linking stress appraisal to performance (Larsson et al. 2007; LeBlanc 2009; Arora et al. 2010; Larsson and Sanner 2010; McGrath et al. 2011).

Aims

The present study is a response to the increasing recognition of the influence of challenge and threat evaluations in a range of medical disciplines and to calls for ongoing

research into the relationship between stress and performance (Larsson et al. 2007; LeBlanc 2009; Arora et al. 2010; Larsson and Sanner 2010; McGrath et al. 2011). Specifically, we aimed to investigate, in the context of selection to specialty training programmes in anaesthesia, (a) which selection centre and candidate factors were associated with variation in the perceived challenge/threat to candidates, and (b) whether, after accounting for these factors, the challenge/threat state was a predictor of candidate performance in the selection centre stations.

Methods

Study context and ethical procedure

The study formed part of a larger project investigating selection into specialty training that was included in the UK's National Institute for Health Research Clinical Research Network (NIHR CRN) Portfolio (Study ID: 9732). The Chair of the NHS Cornwall and Plymouth Research Ethics Committee, UK reviewed the project proposal and confirmed that formal ethical submission was not required. Study participants were doctors applying for specialty training posts in anaesthesia at the South West Peninsula Deanery and who attended one of three selection centres: two for doctors applying at initial-entry (CT1) grade and one for partially trained anaesthetists applying at higher (ST3) grade. Research was carried out according to American Psychological Association ethical guidelines for research with human subjects (American Psychological Association 2002). Study participants were briefed by study investigators, provided with written information regarding data confidentiality and the right to withdraw at any time during the study, and consented to the use of their selection centre data by the study team. This data included both selection centre performance scores and the candidates' demographic and professional profiles, consisting of their age (years), gender, ethnicity (white British or not), country of training (UK or elsewhere), time since registration (months) and portfolio self-score. At the point of application candidates scored the content of their own portfolio in eleven domains (Table 1), covering qualifications and awards during both undergraduate and postgraduate training, anaesthesia-specific experience, research participation and outputs, and teaching involvement. The scoring process comprised detailed descriptors identifying levels of achievement at each available scale point (Appendix 1). These descriptors are essentially 'factual' in nature though there is room for differences in interpretation in light of the widely varying nature of candidates' education histories and career paths. Selection centre assessors were blinded as to whether candidates had consented to take part in the study.

Selection centres

The selection centres, which were conducted as part of the Royal College of Anaesthetists' national selection process, incorporated either five (CT1) or four (ST3) stations. The stations comprised: a structured interview, a portfolio review, a presentation, a mannequin-based teamwork simulation and a telephone communication station. The first three of these form the core of the current Royal College of Anaesthetists' national selection process and, along with the simulation station, have been previously described in detail (Gale et al. 2010). As part of the portfolio review station the assessors ratified or adjusted the candidate's portfolio self-score and this assessor score then contributed to the candidate's selection centre total. The telephone station, which ST3 candidates did not undergo, required candidates to relay important clinical details to an anaesthetic colleague regarding a patient requiring emergency surgery. In each station two consultant assessors independently scored candidates on three non-technical skills and 'overall performance' using 5-point scales

(1=Unsatisfactory; 2=Weak; 3=Typical; 4=Very good; 5=Outstanding). Station scores were a weighted sum of the ratings from each assessor and could range from 10 to 50 points. Candidates rotated through the stations in different orders, which we recorded for each candidate.

Data collection

We asked candidates' to make demand / resource evaluations for each station, immediately prior to entering the room, but after reading information that described the upcoming station. We used two items: '*How demanding do you expect the station to be?*' and '*How able are you to cope with the station?*' both measured on a six point scale (1=Not at all; 6=Extremely). Dividing the score on the first item (demands) by the score on the second (resources) produced a measure of the candidate's challenge / threat state: the Cognitive Appraisal Ratio (CAR) (Tomaka et al. 1993). Values of the CAR greater than one indicate that demands outweigh resources and so correspond to tasks presenting a threat to the candidate, whereas values less than one correspond to challenge evaluations where the candidate perceives their resources as being more than sufficient to meet the demands of the task (Moore et al. 2012; Moore et al. 2013). In line with other studies that have used the CAR for stress-based research, we gave the participants brief instructions in how to complete the two items in the CAR but supplied no definitions of what is meant by either "demanding" or the ability to "cope".

Statistical Analysis

To determine the objectivity of the candidates' portfolio self-scores we correlated them with the assessor-awarded portfolio scores from the portfolio station. We described the demographic profile of the participant sample and the distributions of the CARs recorded prior to each individual station in the assessment centre. In order to investigate the stability of participants' threat/challenge assessments across stations we calculated correlations between the CARs. We conducted two mixed model linear regression analyses, both incorporating a random effect for candidate to account for the clustering in the data. The first used the CAR as the outcome variable with the nature of the station, the order in which the candidate encountered the station (an integer from 1 to 5) and the candidates' demographic and professional profile measures (described above) as predictors. The second model used the candidate's station performance score as the outcome variable and the CAR, nature of the station, order of encountering the station and candidates' demographic and professional profile measures, as predictors. The proportion of station to station variance in performance scores accounted for by the CAR was estimated using the method of Raudenbush and Bryk(Raudenbush and Bryk 2002).

The order of encountering each station was recorded as an integer from 1 to 5 and entered as a covariate in each model to capture any possible linear increase in the outcome variable as the candidate progressed through the stations. To investigate alternative approaches to capturing this order effect we re-ran each model with (i) the order variable entered as a categorical factor, and (ii) the order variable replaced by the identity of the previous station. Results for these alternative models were essentially the same as for the initial model and so are not reported here. Analyses were conducted in IBM SPSS version 20.

Results

The 94 participants comprised 37 and 28 doctors respectively who attended the first and second CT1 selection centres, and 29 who attended the ST3 selection centre. No doctors declined participation. Just over half (54%) of the doctors were male, 94% were trained in the

UK and 83% reported their ethnicity as 'white British'. The median age at the time of interview was 28 years and the median period of GMC registration was 18 months. The objectivity of the candidates' portfolio self-scores was confirmed by a correlation of 0.94 with the station assessors' scores.

Overall, candidates perceived the selection centre stations as more of a threat than a challenge: 55% of all station CARs were above one (i.e. perceived demands outweighing perceived resources). The station CARs varied between 0.40 and 4.00 with the mean being greater than one in all five stations and highest in the presentation and simulation stations (Table 2). Eight of the ten correlation coefficients between the station CARs were significantly positive (Table 3).

Predictors of candidates' challenge / threat state

The first regression analysis (Table 4) confirmed that, after allowing for other variables in the model, the nature of the station was indeed a significant predictor of the CAR ($p=0.001$), along with the candidate's gender ($p<0.001$). In particular, the presentation and simulation stations elicited the highest threat states while female candidates reported higher threat states than their male counterparts. Neither the candidate's age, ethnicity, country of training, time since registration, application grade, order of encountering the station nor their portfolio self-score were predictors of the station CARs.

Predictors of candidates' station performance

The second regression analysis (Table 5) showed that candidates' station performance scores were significantly predicted by their country of medical training, the strength of their portfolio self-assessment, their challenge/threat evaluation of the station (the CAR) and the nature of the station itself. Candidates' gender, age, ethnicity, application grade, time since registration and order of encountering the stations were all unrelated to their performance scores. UK-trained doctors attained higher scores, as did those who presented with stronger portfolio self-scores. The significant variation in scores between the stations reflects the varying difficulty levels of the tasks involved, with scores being highest in the portfolio station and lowest in the simulation station. To check that any possible influence of gender on the station scores was not being masked by the presence of the CAR in the regression model we re-ran the model with the CAR excluded but gender remained non-significant (results not shown).

After controlling for candidates' demographic and professional profiles and for station difficulty and order, the CAR was a significant predictor of station performance. The tendency to evaluate stations as a threat rather than a challenge was associated with poorer performance: a one unit increase in the CAR resulted in an average performance decrease of 2.7 points. While statistically significant this represents a fairly small effect: a one standard deviation (0.46) increase in the CAR corresponded to a 0.15 standard deviation (1.24 point) decrease in the station performance score and the CAR accounted for 3.6% of the variance in those scores.

Discussion

Cognitive appraisal ratios (CARs) for the selection centre stations were related to the content of the station but not to the order in which it was encountered. The simulation and presentation stations induced the strongest threat state and the portfolio review station the lowest. This may not be surprising as the simulation station involved the candidates being observed performing in an unpredictable clinical situation. Similarly, the presentation station

required participants to prepare and deliver a talk on an unprepared subject, a task that bears strong similarity to the Trier Social Stress Test: a frequently used method of inducing stress responses in experimental studies (Kirschbaum et al. 1993). In contrast to the unpredictability inherent in these stations, the portfolio review involves a discussion around documentation that the candidate has prepared and submitted to the panel in advance (Gale et al. 2010). These results are consistent with findings in psychology, which suggest that stronger stress responses result from tasks that are uncontrollable and/or open to observation and potential negative evaluation by others (Dickerson and Kemeny 2004).

An additional driver of the appraisal ratios was the candidate's gender, with female candidates generally reporting higher threat states than their male counterparts. Previous research has also found that women report being more threatened than men prior to undertaking a stressful task, although it is unclear whether this is a real effect or due to gender biases in reporting threat (Quigley et al. 2002). It may simply be that women are less reticent than men in owning up to feeling threatened. Despite differences in perceived threat states however, male and female candidates performed comparably in the selection centre stations. This mirrors research on anxiety in medical students where, despite performing at a level equivalent to their male peers, female medical students consistently report more anxiety and less confidence in their abilities (Blanch et al. 2008). No other demographic or professional profile variables influenced the CARs. Notably, candidates' pre-station threat states were unrelated to their age, time since registration, and past achievements as measured by their portfolio self-scores. This lack of association is both interesting and surprising as we might expect age, experience and previous achievements to strengthen a person's appraisal of their own coping resources. This finding suggests that while threat states may relate to the stress of particular situations, a tendency towards either challenge or threat evaluations may be a stable personality characteristic that is little affected by age or experience. Furthermore, the weak to moderate positive correlations observed between the station CARs suggest some degree of consistency in individuals' evaluations across situations. These findings are consonant with other research showing that stress appraisal styles are related to personality factors and that stress appraisals are moderately stable across situations (Power and Hill 2010).

After adjusting for candidate factors and for the nature of each station, CARs were significantly predictive of station performance. Heightened threat states were associated with poorer performance, although the effect size was fairly small. This indicates that individuals who evaluate the selection centre tasks as more of a challenge than a threat will tend to perform better and hence are more likely to be appointed to a training post. This finding mirrors recent research in both medical and non-medical environments that has also found challenge/threat scores to be predictive of performance in stressful and meaningful evaluative settings. For example, medical students who evaluated an upcoming laparascopic training task as more of a challenge (i.e. personal coping resources matching or exceeding task demands) completed the task more quickly and used visual attention strategies more closely aligned to those used by expert surgeons than those who were in a more threat-like state (i.e., task demands exceeding personal coping resources) (Vine et al. 2013). Pottier and colleagues (Pottier et al. 2013) found that higher pre-scenario threat states were associated with poorer diagnostic accuracy in simulated ambulatory consultations, though not with four other measures of performance. These results were consistent across three separate measures of challenge/threat state, the CAR, the STAI (State Trait Anxiety Inventory) (Spielberger 2010) and salivary cortisol levels, underlining the validity of the simple 2-item CAR. Similar findings of poorer performance by those reporting greater threat evaluations have been found

in relation to competitive golf (Moore et al. 2013), and formalised aviation license assessments (Vine et al. 2015).

Despite the encouraging findings, the present study is not without its limitations. Due to concerns with self-report measures (e.g., response bias), some authors have used objective physiological measures of challenge and threat such as changes in cardiovascular response or salivary cortisol (Blascovich et al. 2004; Harvey et al. 2010; Seery 2011). Unfortunately, it was logistically impossible to fit each participant with an impedance cardiograph device before the selection procedure and, since typical salivary cortisol response times (20-40 min) are greater than the duration of the selections stations, the present study instead used a self-report measure that has been widely employed in the literature and shown to closely correlate with both cardiovascular and cortisol-based indices of challenge and threat (Harvey et al. 2010; Zanstra and Johnston 2011). Alternative self-reported psychological measures of challenge and threat such as the STAI, the Stress Appraisal Measure (Peacock and Wong 1990), or the Stressor Appraisal Scale (Schneider 2008) might have been employed but given the high-stakes nature of the selection centres, their considerably greater length than the 2-item CAR would have risked participant non-completion or study withdrawal. While the ecologically valid setting prevented the level of experimental control required to examine potential mechanisms (including psychophysiological markers), this is one of very few studies that have assessed challenge and threat appraisals immediately prior to performance in a naturalistic, meaningful and self-relevant (rather than an experimental) setting (Tomaka et al. 1993), where performance failure has important consequences. Furthermore, the study addresses weaknesses in other studies where CAR was only taken once, before competition commenced (Moore et al. 2013). In contrast, in the current study we assessed CAR throughout the selection process (before each station) enabling the potentially dynamic nature of demand/resource evaluations to be accounted for.

Managing responses to stress is an important non-technical skill and our findings may have implications for both selection and training in anaesthesia and other acute specialties. Performance in the current selection centre process has been shown to correlate with performance during training (Gale et al. 2010). Combined with the present finding that greater challenge evaluations predict stronger selection centre performance, it would be tempting to conclude that those disposed to challenge rather than threat evaluations may perform better in clinical practice but we cannot reach such a conclusion in the present study. There is no evidence that the effect observed in the selection centres would exist outside of the environment in which the data were captured. Our findings do however support the ongoing use of simulated, task-based assessment for selection to specialty training programmes. The simulation station resulted in the highest threat scores and therefore provides a useful indication of how trainees may perform under the type of acute pressure they will experience in clinical practice.

It has been suggested that such information about how stress is evaluated might support recruitment decisions and subsequent training in both anaesthesia and other acute specialties such as surgery (McGrath et al. 2011). However, while we have shown that individual differences are detectable in candidates' performance in the controlled environment of a selection centre, there are potential limitations to using cognitive appraisal data to inform recruitment and management decisions. First, interpretation of CAR data would need to take account of gender differences in reporting threat states; second the CAR was only weakly associated with performance in the selection stations; third, the CAR is a self-report measure and so open to bias and distortion if openly used in a high-stakes assessment. The evidence for personal consistency in evaluations across situations and over time implies that individuals may have a trait-like quality that predisposes them to habitually appraise

situations as challenging or threatening. From a recruitment perspective, this might provide a rationale for including high threat tasks in selection centres for branches of medicine that demand resilience to acutely stressful situations. Anaesthetists work in stressful environments where difficult clinical situations can be construed as either a threat or a challenge. Larsson and Sanner have highlighted the strategies employed by experienced anaesthetists to control their reactions to such situations and have emphasised the danger that anaesthesia trainees could suffer burnout if repeatedly exposed to difficult clinical situations without possessing the coping mechanisms to deal with them (Larsson and Sanner 2010). The authors suggested that cognitive reappraisal may help trainees deal with acutely stressful situations but we have found that trainees vary markedly in the way that they appraise such situations, certainly in the context of dealing with work-related tasks at selection centres. Whether trainees can successfully learn adaptive behaviour in these instances is an area for future research.

From a training perspective, interventions aimed at ensuring that doctors *learn* to evaluate stressful events as a challenge rather than a threat could have clinical utility (Harvey et al. 2010; McGrath et al. 2011). This could be achieved by helping trainees change their perceptions of task demands, or by altering their actual or perceived resources (Larsson et al. 2007; Feinberg and Aiello 2010; Larsson and Sanner 2010). In terms of actual resources available there is an increasing amount of evidence favouring the use of 'decision support tools' such as algorithms and protocols to improve clinicians' performance and decision making in life threatening situations (Skinner and Brewer 2002; Wayne et al. 2005; Harrison et al. 2006; Nelson et al. 2008; Low et al. 2011). Human factors and crisis resource management training aims to promote efficient use of resources in these situations; factors affecting perceived resources which can be targeted include familiarity, uncertainty, difficulty, danger, attitudes, and the presence of others (Seery 2011). Other methods of learning to deal more effectively with stressful situations, which have been tried or suggested for use with healthcare professionals, include 'stress inoculation training' (Meichenbaum 1985), 'mindfulness' (Hayes and Smith 2005) and a range of cognitive behavioural methods (Flin et al. 2008; Norris and Lockey 2012; Regehr et al. 2014). CARs could potentially be used to indicate the individuals for whom such training would be most beneficial.

Conclusions

This study demonstrates that challenge and threat states, assessed just prior to an anaesthesia selection centre station, are statistically significant predictors of performance in that station, though the effect size is small. Specifically, a challenge state was associated with superior station performance compared to a threat state. Stations incorporating work related tasks induced higher threat states in applicants and may be able to discern how future anaesthetists will react under similar situations in the clinical environment. In anaesthesia, as in other acute specialties, performing effectively under heightened stress is a necessity and it is important therefore, that recruitment processes can identify doctors who possess the appropriate cognitive resources to deal with challenging situations. This study suggests that our multi-station selection centres may help in achieving that goal but the results are not strong enough to support the direct use of cognitive appraisal data in recruitment decisions.

Acknowledgements

We would like to acknowledge the assistance of the staff of the South West Peninsula Deanery, Plymouth, UK, staff from the Horizon Centre, Torbay Hospital, Torquay, UK, and healthcare professionals from hospitals in the region who were involved in running the selection centres. We also thank Sam Vine, University of Exeter, for valuable comments on earlier drafts of this paper and for information regarding his research on stress evaluations in airline pilots, and Vicki LeBlanc, University of Toronto, for perceptive feedback on a near-final version of the paper.

The Royal College of Anaesthetists kindly gave permission to reproduce the detailed descriptors used to score candidates' portfolios of academic and professional achievement (Appendix 1).

References

American Psychological Association (2002). Ethical Principles of Psychologists and Code of Conduct. *American Psychologist, 57*(12), 1060-1073.

Arora, S., Sevdalis, N., Nestel, D., Woloshynowych, M., Darzi, A., & Kneebone, R. (2010). The impact of stress on surgical performance: A systematic review of the literature. *Surgery, 147*(3), 318-330.e6.

Blanch, D. C., Hall, J. A., Roter, D. L., & Frankel, R. M. (2008). Medical student gender and issues of confidence. *Patient Education and Counseling, 72*(3), 374-381.

Blascovich, J. (2008). Challenge and threat. In A. J. Elliot, *Handbook of approach and avoidance motivation*. New York: Psychology Press, 431-445.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. P. Forgas, *Feeling and thinking: The role of affect in social cognition*. Cambridge: Cambridge University Press, 59-82.

Blascovich, J., Seery, M. D., Mugridge, C. A., Norris, R. K., & Weisbuch, M. (2004). Predicting athletic performance from cardiovascular indexes of challenge and threat. *Journal of Experimental Social Psychology, 40*(5), 683-688.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin, 130*(3), 355-391.

Feinberg, J. M., & Aiello, J. R. (2010). The Effect of Challenge and Threat Appraisals Under Evaluative Presence. *Journal of Applied Social Psychology, 40*(8), 2071-2104.

Flin, R. H., O'Connor, P., & Crichton, M. (2008). *Safety at the sharp end: a guide to non-technical skills*. Farnham: Ashgate.

Gale, T. C. E., Roberts, M. J., Sice, P. J., Langton, J. A., Patterson, F. C., Carr, A. S., et al. (2010). Predictive validity of a new selection centre testing non-technical skills for recruitment to training in anaesthesia. *British Journal of Anaesthesia, 105*(5), 603-609.

Harrison, T. K., Manser, T., Howard, S. K., & Gaba, D. M. (2006). Use of Cognitive Aids in a Simulated Anesthetic Crisis. *Anesthesia and Analgesia, 103*(3), 551-556.

Harvey, A., Nathens, A. B., Bandiera, G., & LeBlanc, V. R. (2010). Threat and challenge: cognitive appraisal and stress responses in simulated trauma resuscitations. *Medical Education, 44*(6), 587-594.

Hayes, S. C., & Smith, S. (2005). *Get out of your mind and into your life*. Oakland, CA: New Hardinger Publications.

Hunziker, S., Semmer, N. K., Tschan, F., Schuetz, P., Mueller, B., & Marsch, S. (2012). Dynamics and association of different acute stress markers with performance during a simulated resuscitation. *Resuscitation, 83*(5), 572-578.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test' – A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology, 28*(1-2), 76-81.

Larsson, J., Rosenqvist, U., & Holmström, I. (2007). Enjoying work or burdened by it? How anaesthetists experience and handle difficulties at work: a qualitative study. *British Journal of Anaesthesia, 99*(4), 493-499.

Larsson, J., & Sanner, M. (2010). Doing a good job and getting something good out of it: on stress and well-being in anaesthesia. *British Journal of Anaesthesia, 105*(1), 34-37.

LeBlanc, V. R. (2009). The Effects of Acute Stress on Performance: Implications for Health Professions Education. *Academic Medicine, 84*(10), S25-S33.

Low, D., Clark, N., Soar, J., Padkin, A., Stoneham, A., Perkins, G. D., et al. (2011). A randomised control trial to determine if use of the iResus© application on a smart

phone improves the performance of an advanced life support provider in a simulated medical emergency. *Anaesthesia, 66*(4), 255-262.
McGrath, J. S., Moore, L., Wilson, M. R., Freeman, P., & Vine, S. (2011). 'Challenge' and 'threat' states in surgery: implications for surgical performance and training. *BJU International, 108*(6), 795-796.
Meichenbaum, D. (1985). *Stress inoculation training*. Oxford: Pergamon Press.
Moore, L. J., Vine, S. J., Wilson, M. R., & Freeman, P. (2012). The effect of challenge and threat states on performance: An examination of potential mechanisms. *Psychophysiology, 49*(10), 1417-1425.
Moore, L. J., Wilson, M. R., Vine, S. J., Coussens, A., & Freeman, P. (2013). Champ or chump? Challenge and threat states during pressurized competition. *Journal of Sport and Exercise Psychology, 35*, 551-562.
Nelson, K. L., Shilkofski, N. A., Haggerty, J. A., Saliski, M., & Hunt, E. A. (2008). The use of cognitive AIDS during simulated pediatric cardiopulmonary arrests. *Simulation in healthcare, 3*(3), 138-145.
Norris, E. M., & Lockey, A. S. (2012). Human factors in resuscitation teaching. *Resuscitation, 83*(4), 423-427.
Peacock, E. J., & Wong, P. T. P. (1990). The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal. *Stress Medicine, 6*(3), 227-236.
Piquette, D., Tarshis, J., Sinuff, T., Fowler, R. A., Pinto, R., & Leblanc, V. R. (2014). Impact of Acute Stress on Resident Performance During Simulated Resuscitation Episodes: A Prospective Randomized Cross-Over Study. *Teaching and Learning in Medicine, 26*(1), 9-16.
Pottier, P., Dejoie, T., Hardouin, J.-B., Le Loupp, A.-G., Planchon, B., Bonnaud, A., et al. (2013). Effect of stress on clinical reasoning during simulated ambulatory consultations. *Medical Teacher, 35*(6), 472-480.
Pottier, P., Hardouin, J.-B., Dejoie, T., Bonnaud, A., Le Loupp, A.-G., Planchon, B., et al. (2011). Stress responses in medical students in ambulatory and in-hospital patient consultations. *Medical Education, 45*(7), 678-687.
Power, T. G., & Hill, L. G. (2010). Individual differences in appraisal of minor, potentially stressful events: A cluster analytic approach. *Cognition & Emotion, 24*(7), 1081-1094.
Quigley, K. S., Barrett, L. F., & Weinstein, S. (2002). Cardiovascular patterns associated with threat and challenge appraisals: A within-subjects analysis. *Psychophysiology, 39*(3), 292-302.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks: Sage.
Regehr, C., Glancy, D. B., Pitts, A., & LeBlanc, V. R. (2014). Interventions to Reduce the Consequences of Stress in Physicians: A Review and Meta-Analysis. *Journal of Nervous and Mental Disease, 202*(5), 353-359.
Schneider, T. R. (2008). Evaluations of stressful transactions: what's in an appraisal? *Stress and Health, 24*(2), 151-158.
Seery, M. D. (2011). Challenge or threat? Cardiovascular indexes of resilience and vulnerability to potential stress in humans. *Neuroscience and Biobehavioral Reviews, 35*(7), 1603-1610.
Skinner, N., & Brewer, N. (2002). The dynamics of threat and challenge appraisals prior to stressful achievement events. *Journal of Personality and Social Psychology, 83*, 678-692.
Spielberger, C. D. (2010). State-Trait Anxiety Inventory. In *The Corsini Encyclopedia of Psychology*: John Wiley & Sons.

Tomaka, J., Blascovich, J., Kelsey, R. M., & Leitten, C. L. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology, 65*(2), 248-260.

Vine, S. J., Freeman, P., Moore, L. J., Chandra-Ramanan, R., & Wilson, M. R. (2013). Evaluating stress as a challenge is associated with superior attentional control and motor skill performance: Testing the predictions of the biopsychosocial model of challenge and threat. *Journal of Experimental Psychology: Applied, 19*(3), 185-194.

Vine, S. J., Uiga, L., Lavric, A., Moore, L. J., Tsaneva-Atanasova, K., & Wilson, M. R. (2015). Individual reactions to stress predict performance during a critical aviation incident. *Anxiety, Stress, & Coping, 28*(4), 467-477.

Wayne, D. B., Butter, J., Siddall, V. J., Fudala, M. J., Linquist, L. A., Feinglass, J., et al. (2005). Simulation-Based Training of Internal Medicine Residents in Advanced Cardiac Life Support Protocols: A Randomized Trial. *Teaching and Learning in Medicine, 17*(3), 202-208.

Zanstra, Y. J., & Johnston, D. W. (2011). Cardiovascular reactivity in real life settings: Measurement, mechanisms and meaning. *Biological Psychology, 86*(2), 98-105.

Tables

Table 1: Domains for assessment of portfolio content and maximum score by year and grade of application.

	Application grade	
Domain	CT1	ST3
UG Training - Additional Degrees	6	6
UG Training - Prizes and Awards	2	2
PG Medical Qualifications	4	4
PG Training in other specialties	4	4
Experience (CT1) / Qualifications (ST3) in Anaesthesia	6	5
Clinical Governance and Audit	5	5
Research	5	6
Teaching	5	5
Academic Publications	5	5
Presentations and Posters	4	4
Training Courses attended	4	4
Total	50	50

UG = Undergraduate, PG = Postgraduate

Table 2: Descriptive statistics for cognitive appraisal ratios by selection centre station

Station	N	Minimum	Maximum	Mean	SD	Skewness
Interview	91	0.60	2.50	1.22	0.30	1.33
Portfolio	92	0.40	4.00	1.16	0.52	2.43
Presentation	93	0.40	3.00	1.33	0.55	1.40
Simulation	89	0.50	3.00	1.34	0.47	1.45
Telephone	64	0.60	2.50	1.18	0.35	1.15
All	429	0.40	4.00	1.25	0.46	1.78

Table 3: Correlations between the cognitive appraisal ratios

Station	Portfolio	Presentation	Simulation	Telephone
Interview	0.35**	0.52**	0.39**	0.38**
Portfolio		0.35**	0.07	0.06
Presentation			0.24*	0.28*
Simulation				0.49**

* Correlation significant at the 0.05 level (2-tailed).
** Correlation significant at the 0.01 level (2-tailed).

Table 4: Regression of cognitive appraisal ratio (CAR) on candidate demographic profile, portfolio self-score, station and station order.

	Coefficient	Std. Error	t	p	LCL	UCL
Female	0.229	0.060	3.834	0.000	0.110	0.348
White British	0.048	0.094	0.508	0.613	-0.140	0.236
Age	0.020	0.011	1.837	0.070	-0.002	0.041
CT1 application grade	0.002	0.095	0.022	0.983	-0.186	0.190
UK-trained	0.116	0.149	0.773	0.441	-0.181	0.412
Months registered	-0.003	0.002	-1.416	0.160	-0.007	0.001
Portfolio self-score	-0.005	0.004	-1.103	0.273	-0.013	0.004
Station (Telephone = ref. category)*						
Interview	0.072	0.065	1.096	0.274	-0.057	0.200
Portfolio	0.006	0.066	0.092	0.927	-0.123	0.135
Presentation	0.181	0.065	2.776	0.006	0.053	0.308
Simulation	0.190	0.065	2.918	0.004	0.062	0.319
Station order	-0.004	0.014	-0.303	0.762	-0.032	0.024

* F-test for overall significance of 'Station' : $F_{334,4}$=4.61, p=0.001

Table 5: Regression of station performance score on candidate characteristics, station, station order and station cognitive appraisal ratio (CAR).

	Coefficient	Std. Error	t	p	LCL	UCL
Female	-0.004	1.028	-0.004	0.997	-2.046	2.039
White British	0.079	1.593	0.050	0.960	-3.089	3.248
Age	-0.167	0.183	-0.911	0.365	-0.532	0.198
CT1 application grade	-0.838	1.602	-0.523	0.602	-4.018	2.342
UK-trained	10.906	2.523	4.322	0.000	5.891	15.920
Months registered	0.009	0.033	0.263	0.793	-0.056	0.074
Portfolio self-score	0.151	0.070	2.138	0.035	0.011	0.290
Station (Telephone = ref. category)*						
Interview	-1.229	1.136	-1.082	0.280	-3.464	1.005
Portfolio	1.531	1.141	1.342	0.180	-0.713	3.776
Presentation	1.094	1.141	0.959	0.338	-1.150	3.338
Simulation	-4.638	1.145	-4.049	0.000	-6.891	-2.385
Station order	0.324	0.249	1.303	0.193	-0.165	0.813
Cognitive appraisal ratio (CAR)	-2.738	0.849	-3.225	0.001	-4.408	-1.069

* F-test for overall significance of 'Station' : $F_{335,4}$=11.68, p<0.001

Appendix 1: Criteria for assessment of candidates' portfolios

The following tables give the detailed criteria that were used to assess candidates' portfolios of academic and professional achievement. These criteria were used both by candidates, to self-assess their own portfolios at the point of application, and by assessors (who verified evidence presented by candidates at the portfolio station of the selection centre). The tables are reproduced here by kind permission of the Royal College of Anaesthetists.

Table A1: Portfolio assessment criteria, CT1 grade.

Portfolio Assessment Criteria
Anaesthesia CT1, 2013

Domain	Descriptor	Score
Undergraduate Training - Additional Degrees	First-class intercalated BSc (or BA or similar/higher degree) OR First-class BSc (or similar/higher) prior to starting medicine and relevant to medicine/anaesthesia	6
	First-class BA (or similar/higher) prior to starting medicine and not relevant to medicine/anaesthesia	5
	2.1 intercalated BSc (or BA or similar/higher degree) OR 2.1 BSc (or similar/higher) prior to starting medicine and relevant to medicine/anaesthesia	4
	2.1 BA (or similar/higher) prior to starting medicine and not relevant to medicine/anaesthesia	3
	2.2 or third class intercalated BSc (or BA or similar/higher) OR 2.2/third class BSc (or similar/higher) prior to starting medicine and relevant to medicine/anaesthesia	0
	MD linked to Primary Medical qualification	0
	None of the above	0

Portfolio Assessment Criteria
Anaesthesia CT1, 2013

Domain	Descriptor	Score
Undergraduate Training - Prizes and Awards	I was awarded a Distinction during my undergraduate medical training	2
	I was awarded a Prize or Merit during my undergraduate medical training	1
	I was not awarded one	0
Postgraduate Medical Qualifications (excluding anaesthesia)	UK final postgraduate qualification e.g. MRCP	4
	UK primary postgraduate qualification e.g. MRCP Part 1, MCEM A	2
	I do not have one	0
Postgraduate Training in other specialties (excluding Foundation)	19 to 24 months experience in an appropriate/relevant Complementary Specialty (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by the time of appointment, with achievement of expected outcomes, anywhere in the world	4
	13 to 18 months experience in an appropriate/relevant Complementary Specialty (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by the time of appointment, with achievement of expected outcomes, anywhere in the world	2
	More than 24 months in Complementary Specialties (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), with achievement of expected outcomes, anywhere in the world	2
	5 to 12 months experience in an appropriate/relevant Complementary Specialty (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by the time of appointment, anywhere in the world	1
	Experience in Complementary Specialties (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), without achievement of all expected outcomes, anywhere in the world	1

Portfolio Assessment Criteria
Anaesthesia CT1, 2013

Domain	Descriptor	Score
	No experience in a Complementary specialty (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology)	0
Experience in anaesthesia	Post foundation experience in Anaesthesia/ICM (less than 18 months) and 2 or more of the following – foundation post in Anaesthesia/ICM, undergraduate Special Study module, undergraduate Student Selected component, undergraduate elective, undergraduate taster session	6
	Post foundation experience in Anaesthesia/ICM (less than 18 months)	5
	2 or more of the following complete by appointment – foundation post in Anaesthesia/ICM, undergraduate Special Study module, undergraduate Student Selected component, undergraduate elective, undergraduate taster session	4
	1 of the following complete by appointment – foundation post in Anaesthesia/ICM, undergraduate Special Study module, undergraduate Student Selected component, undergraduate elective, undergraduate taster session	2
	None of the above	0
Clinical Governance and Audit	Have made a significant contribution to postgraduate audit projects (e.g. managed an audit project, made recommendations for changes to practice based on audit findings, closed the audit loop) - more than one project a year	5
	Have made a significant contribution to postgraduate audit projects (e.g. managed an audit project, made recommendations for changes to practice based on audit findings, closed the audit loop) - one a year or less	3
	Have made a significant contribution to/instigated an undergraduate audit project	2
	Have taken part in other people's audits (ie data collection) or made other minor contributions	1
	I have not participated in an audit project	0

Portfolio Assessment Criteria
Anaesthesia CT1, 2013

Domain	Descriptor	Score
Research	Completion of a PhD with original research	5
	Significant involvement in research eg two-year MD with original research	4
	I have had some personal and direct involvement with planning or running a postgraduate research project	3
	I have had significant personal involvement with an undergraduate research project	1
	I understand the importance of research	0
Teaching	I have made a major contribution to a local/national teaching programme including organising a programme OR have a recognised qualification in teaching e.g. Diploma in Medical Education, Masters in Medical Education Certificate in Medical Education	5
	I have some experience in formal teaching Health Professionals (eg lectures) and/or I have some experience of formal teaching Health-related topics to a non-medical audience (eg lectures) and/or I have attended a Teaching the Teachers course, Generic Instructors course (or similar)	2
	I have carried out informal teaching of colleagues (including nurses and medical students)	1
	I have not contributed to teaching	0
Academic Publications (to include undergraduate and/or postgraduate)	I am first author in more than one peer-reviewed publication	5
	I am first author on one peer-reviewed publication	4
	I am co-author in more than one peer-reviewed publication	3
	I am co-author in one peer-reviewed publication	2
	I have published one or more abstracts or articles (including e-publications) or I have published a letter	1

Portfolio Assessment Criteria
Anaesthesia CT1, 2013

Domain	Descriptor	Score
	I have not published anything	0
Presentations (including audit presentations)	I have presented at a regional, national or international meeting	4
	I have presented at local/departmental meetings on more than one occasion	2
	I have made a local/departmental presentation on one occasion	1
	I have made no presentations nor shown any posters	0
Training courses attended (and current) at time of interview	3 or more courses attended (including ALS, ATLS, ILS, BASICS, IMPACT, APLS, EPLS, College and Simulator Courses) (or similar)	4
	2 courses attended (including ALS, ATLS, ILS, BASICS, IMPACT, APLS, EPLS, College and Simulator Courses) (or similar)	2
	1 course attended (including ALS, ATLS, ILS, BASICS, IMPACT, APLS, EPLS, College and Simulator Courses) (or similar)	1
	No courses attended	0

Table A2: Portfolio assessment criteria, ST3 grade.

Portfolio Assessment Criteria Anaesthesia ST3, 2013		
Domain	**Descriptor**	**Score**
Undergraduate Training - Additional Degrees	First-class intercalated BSc (or BA or similar/higher degree) OR First-class BSc (or similar/higher) prior to starting medicine and relevant to medicine/anaesthesia	6
	First-class BA (or similar/higher) prior to starting medicine and not relevant to medicine/anaesthesia	5
	2.1 intercalated BSc (or BA or similar/higher degree) OR 2.1 BSc (or similar/higher) prior to starting medicine and relevant to medicine/anaesthesia	4
	2.1 BA (or similar/higher) prior to starting medicine and relevant to medicine/anaesthesia	3
	2.2 or third class intercalated BSc (or BA or similar/higher) OR 2.2/third class BSc (or similar/higher) prior to starting medicine and relevant to medicine/anaesthesia	0
	MD linked to Primary Medical qualification	0
	None of the above	0
Undergraduate Training - Prizes and Awards	I was awarded a Distinction during my undergraduate medical training	2
	I was awarded a Prize or Merit during my undergraduate medical training	1
	I was not awarded one	0
Postgraduate Medical Qualifications	UK final postgraduate qualification e.g. MRCP, MRCS, MCEM	4
	UK primary postgraduate qualification e.g. MRCP Part 1, MCEM A	2

Portfolio Assessment Criteria
Anaesthesia ST3, 2013

Domain	Descriptor	Score
(excluding anaesthesia and ICM)	None of the above	0
Postgraduate Training in other specialties (including research posts but excluding Foundation, Anaesthesia or ICM)[1]	19 to 24 months experience in an appropriate/relevant Complementary Specialty (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by the time of appointment, with achievement of expected outcomes, anywhere in the world	4
	13 to 18 months experience in an appropriate/relevant Complementary Specialty (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by the time of appointment, with achievement of expected outcomes, anywhere in the world	2
	More than 24 months in Complementary Specialties (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by time of appointment, with achievement of expected outcomes, anywhere in the world	2
	5 to 12 months experience in an appropriate/relevant Complementary Specialty (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by the time of appointment, anywhere in the world	1
	More than 24 months in Complementary Specialties (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by time of appointment, without achievement of all expected outcomes, anywhere in the world	1
	No experience in a Complementary specialty (e.g. Surgery, Medicine, Emergency Medicine, Paediatrics, Obstetrics, Radiology), by time of appointment	0
Qualifications in Anaesthesia (applicants with	Primary FRCA passed within first year of entering specialty (excluding time in Foundation Programme)	5
	Primary FRCA (or equivalent) all sections (MCQ and OSCE/viva) passed at first attempt	4

[1] Experience gained as general duties medical officer with the Defence Medical Services will not be considered in this domain.

Portfolio Assessment Criteria
Anaesthesia ST3, 2013

Domain	Descriptor	Score
Final FRCA should score depending on success at Primary examination attempts only)	Primary FRCA (or equivalent) one section (MCQ or OSCE/viva) passed at first sitting (other part passed at a repeat attempt)	3
	Primary FRCA (or equivalent) with MCQ and OSCE/viva both passed at a repeat attempt	2
	I have no qualifications in anaesthesia	0
Clinical Governance and Audit as a postgraduate	Have made a significant contribution to postgraduate audit projects (e.g. managed an audit project, made recommendations for changes to practice based on audit findings, closed the audit loop) - more than one project a year and at least 1 presented at a regional, national or international meeting	5
	Have made a significant contribution to postgraduate audit projects (e.g. managed an audit project, made recommendations for changes to practice based on audit findings, closed the audit loop) - one a year or less and at least 1 presented at a regional National or International meeting	4
	Have made a significant contribution to postgraduate audit projects (e.g. managed an audit project, made recommendations for changes to practice based on audit findings, closed the audit loop) - one a year or less and these presented at a Local meeting	3
	Have taken part in other people's postgraduate audits, (ie data collection) or made other minor contributions, not led or designed the audit	2
	I have not participated in an audit project	0
Research	Completion of a PhD with original research	6
	Significant involvement in research e.g. two-year MD with original research	4
	I have had some personal and direct involvement with planning or running a postgraduate research project	3

Portfolio Assessment Criteria
Anaesthesia ST3, 2013

Domain	Descriptor	Score
	I understand the importance of research	0
Teaching	I have made a major contribution to a local/national teaching programme including organising a programme OR have a recognised qualification in teaching e.g. Diploma in Medical Education, Masters in Medical Education	5
	I have some experience in formal teaching Health Professionals (e.g. lectures) and/or I have some experience of formal teaching Health-related topics to a non-medical audience (e.g. lectures) and/or I have attended a Teaching the Teachers course, Generic Instructors course or similar	2
	I have carried out informal teaching of colleagues (including nurses and medical students)	1
	I have not contributed to teaching	0
Academic Publications (undergraduate or postgraduate)	I am first author in more than one peer-reviewed publication as a postgraduate	5
	I am first author on one peer-reviewed publication as a postgraduate	4
	I am co-author in more than one peer-reviewed publication as a postgraduate	3
	I am co-author in one peer-reviewed publication as a postgraduate or I am an author of an undergraduate publication	2
	I have published one or more abstracts or articles (including e-publications) or I have published a letter	1
	I have not published anything	0
Presentations and Poster Presentations (excluding audit	I have presented at a regional, national or international meeting	4
	I have presented at a local meeting on more than one occasion	2
	I have made a local/departmental presentation	1

Portfolio Assessment Criteria
Anaesthesia ST3, 2013

Domain	Descriptor	Score
presentations)	I have made no presentations nor shown any posters	0
Training courses attended (and current) at time of interview	3 or more courses attended (including ALS, ATLS, ILS, BASICS, IMPACT, APLS, EPLS, College and Simulator Courses) (or similar)	4
	2 courses attended (including ALS, ATLS, ILS, BASICS, IMPACT, APLS, EPLS, College and Simulator Courses) (or similar)	2
	I course attended (including ALS, ATLS, ILS, BASICS, IMPACT, APLS, EPLS, College and Simulator Courses) (or similar)	1
	No courses attended	0