

The relationship between acute stress and clinical performance in medical students: a pilot simulation study

Russ SJ¹, Morrison I¹, Bell C², Morse J¹, MacKenzie R¹ & Johnston M²

¹Institute of Education in Medical and Dental Sciences, University of Aberdeen

²Institute of Applied Health Sciences, University of Aberdeen

Corresponding Author: Dr Stephanie J Russ

Corresponding Address: Institute for Education in Medical and Dental Sciences

University of Aberdeen

Polwarth Building Room 2.036,

Foresterhill

Aberdeen

AB25 2ZD

stephanie.russ@abdn.ac.uk

Tel: 07725560505

Word Count: 4182

ABSTRACT

Background: Acute stress has been linked to impaired clinical performance in healthcare settings. However, few studies have measured experienced stress and performance simultaneously using robust measures in controlled experimental conditions, which limits the strength of their findings.

Aim: In the current study we examined the relationship between stress and clinical performance in 2nd year medical students undertaking a simulated ECG scenario. To explore this relationship in greater depth we manipulated two variables (clinical urgency and cognitive load), and also examined the impact of trait anxiety and task self-efficacy.

Methods: 2nd year medical students were asked to conduct a 12-lead ECG on a simulated patient. Students were randomly assigned to one of four experimental conditions according to clinical urgency (high/low) and cognitive load (high/low), which were manipulated during a handover prior to the ECG. During the scenario they were asked to describe the ECG trace to a senior doctor over the phone and to conduct a drug calculation. They then received a performance debrief. Psychological stress and physiological stress were captured (via self-report and heart-rate respectively) and various aspects of performance were observed, including technical competence, quality of communication, work-rate and compliance with patient safety checks. Trait anxiety and task self-efficacy were also captured via self-report.

Results: 50 students participated. While there was little impact of experimental condition on stress or performance, there was a significant relationship between stress and performance for the group as a whole. Technical competence was poorer for those reporting higher levels of psychological stress prior to and following the procedure. Neither trait-anxiety nor task self-efficacy mediated this relationship.

Conclusions: This study has provided evidence for a link between acute stress and impaired technical performance in medical students completing a simulation scenario, using real-time measures. The implications for patient safety and medical education are discussed.

WHAT THIS PAPER ADDS:

What is already known?

- We know that high levels of acute stress can impair cognitive function and performance; however, the impact of stress on clinical performance has been underexplored.
- This is largely due to the methodological challenges inherent in measuring both stress and performance in healthcare settings in a way that makes it possible to link the two. Most have used retrospective approaches with proxy or self-report measures of performance.

What this study adds

- Using simultaneous and robust measures of stress and performance in a simulation setting, this pilot study supports a link between high levels of acute stress and impaired technical performance of clinical skills in medical students.
- In piloting the study we were able to explore the feasibility of the design prior to testing with a larger cohort. As a result, some small changes will be made to the manipulation of cognitive load going forward.
- This study adds strength to the somewhat limited evidence-base in this area and has important implications for patient safety. These findings will be explored in greater depth with future cohorts of medical students.

INTRODUCTION

Hospital doctors typically experience a wide range of stressors in their work setting and report high levels of occupational stress and burnout [1-4]. Clinical work presents a range of stressors which might affect performance – some specific to the clinical situation, e.g. dealing with ill patients, performing clinical tasks, being observed by senior clinicians, alarms, bleeps, telephone, distractions – others are generic, i.e. would constitute stressors in any environment – e.g. cognitive stress of retaining and using large quantities of information, heavy workload, and inadequate support/supervision [5-8]. The experience of occupational stress seems to be particularly marked in junior doctors making the transition from the comparatively protected learning environment of the medical student to the work environment [6, 9-10].

While low intensity stressors can facilitate performance, retention of information and motivation to complete challenging tasks [11], prolonged or excessive levels of stressors that exceed an individual's perceived coping resources can negatively affect cognitive functioning and performance [12-14]. This presents risk for patient safety. In the presence of stressors and associated burnout there is evidence of less than optimal clinical performance, including impaired psychomotor performance, prescribing errors, increased probability of involvement in an adverse event, increased prevalence of malpractice suits, and impaired non-technical skills such as communication and decision making [15-21].

One theory for these adverse impacts is that cognitive load is increased under stressful conditions and puts pressure on limited working memory capacity [22-24]. Further detrimental effects of work stress could be attentional narrowing and distraction. These altered cognitive processes adversely affect professional judgments, decision making, efficiency and motivation or satisfaction, and consequently render the individual more error prone [19]. It has also been postulated that more stable personality traits might moderate these relationships, namely trait anxiety and self-efficacy which have been linked to increased stress and exacerbation of the adverse impact of stress on performance [25-26].

In healthcare, few studies have simultaneously measured stressors and performance due to methodological challenges, making it difficult to draw links between the two. And we are yet to understand fully which conditions are the most stress-inducing and which tasks are most susceptible to disruption by stress. The simulation environment allows reproducible manipulation of variables and observation of real-time impact on performance making it possible to explore these relationships in more depth.

The aim of this pilot study was to measure the effect of stressors on various aspects of clinical performance, including calculations for prescribing, in second year medical students participating in a simulated clinical scenario. (Second year medical students were selected as this was part of an effort to incorporate simulation-based education earlier on in the undergraduate medical curriculum). We

manipulated two variables via experimental conditions – firstly clinical urgency, by varying the severity of the patients’ illness (and thus varying the degree of demand experienced), and secondly cognitive load, by varying the required level of communication with the patient (and thus varying the burden on working memory).

The specific research questions were as follows:

- What is the relationship between stress and clinical performance?
- How does clinical urgency and cognitive load influence these relationships?
- Does trait anxiety and self-efficacy mediate these relationships?

We present data from the first cohort to undertake this simulated learning opportunity. We treated this as a pilot study to test the feasibility of the design before rolling it out with a larger cohort of medical students.

METHOD

Participants

All 180 year 2 medical students in academic year 2014-15 at the University of Aberdeen were invited via email to take part in a simulated clinical scenario to practice their newly acquired ECG and patient management skills. They were advised that they would also be invited to take part in a research study looking at how students react in these settings. At this stage attendance was not mandatory.

Design

This was a cross-sectional, experimental factorial design, with two factors: clinical urgency and cognitive load, each with two levels (high/low). Both factors were manipulated during a handover given to the student immediately before they entered the simulation room to conduct the ECG. Clinical urgency was manipulated by framing the scenario as either clinically urgent (high urgency) or routine (low urgency), while cognitive load was manipulated by asking the student to either explain to the patient precisely what they were doing as they went along (high cognitive load) or not (low cognitive load). Students were randomly allocated to one of four groups according to these factors:

- Group 1: Low clinical urgency, low cognitive load
- Group 2: High clinical urgency, low cognitive load
- Group 3: Low clinical urgency, high cognitive load
- Group 4: High clinical urgency, high cognitive load

The outcome variables were performance and stress. Additionally we assessed task self-efficacy and trait anxiety. Ethical approval for study was granted by the University of Aberdeen College Ethics Review Board.

Measures

Performance

Various clinically relevant performance measures were recorded (Fig 1). These included important patient safety checks (x 4), technical competence assessments (x14), and quality of communication variables (x7), which were all scored by a clinician observer during the scenario (for the purposes of the study and to aid the debrief) and checked by a blinded observer retrospectively from video-recordings. These were all binary variables (yes=1, no=0) but were additionally summed to provide a total patient safety, technical competence and quality of communication score, and an overall total performance score out of 25. We also included two continuous measures of ‘work rate’ (time taken to complete the ECG and a drug calculation) – which were scored retrospectively from the video recordings. Finally, with regards communication, we looked at the percentage of time spent talking during the ECG and the nature of any social communication driven by the students i.e. whether the student initiated conversation about non-clinical topics (e.g. asks about patient’s job). The latter variable was scored on a binary (yes/no) scale but was not included in the ‘total performance score’ as it was decided that social communication could either be positive (e.g. putting the patient at ease) or negative (e.g. non-timely, inappropriate). These measures were informed by OSCE items for an ECG station, but were more comprehensive than the performance variables that would be scored during an OSCE.

Stress

Physiological Stress: Heart rate was captured during the simulation using a small, lightweight Actiheart (CamNTEch) monitor, which measures activity from a chest mounted accelerometer, heart rate from ECG, and computes a measure of energy expenditure by combining the two. Non-metabolically determined changes in heart rate provide a physiological indicator of stress. Students taking medication known to alter heart rate (e.g. beta blockers) were excluded from this element of the study.

Psychological stress: Self-reported anxiety was captured using the shortened version of the State section of the validated State Trait Anxiety Inventory (STAI) [27]. This consists of 6-items which capture feelings pertaining to anxiety at a certain point in time e.g. ‘I feel worried’ ‘I feel tense’. Items are scored using a Likert response scale ranging from 1 (not at all) to 4 (very much so). Total scores

range from 6-24. State anxiety scores vary reliably with situational stress and as such the State scale is commonly used as a proxy measure of acute stress in research studies [28].

Trait anxiety and task self-efficacy

Trait anxiety: Trait anxiety was measured using the Trait section of the STAI. This consists of 20 items designed to capture a stable propensity to experience anxiety, depression and general negative affect, and a tendency to interpret stressful situations as threatening [29]. Items are scored on a 4 point Likert scale ranging from 1 (not at all) to 4 (very much) e.g. 'I am a steady person' 'I feel inadequate'. Total scores range from 20-80.

Self-efficacy: A 9-item self-efficacy scale was constructed for the purposes of this study using published guidance [30]. Items captured self-reported confidence in completing tasks relevant to the scenario e.g. How confident are you that you can ...'perform an ECG without the support of a senior', 'interact with patients' 'perform calculations without a calculator'. Items were scored on a scale ranging from 0-100% confident. Participants indicated where they fell on the scale by reporting a % (e.g. 60%) for each item individually. Total scores on this scale could therefore range from 0-900%, with higher scores indicating greater confidence.

Session Evaluation

A short 6-item self-report scale was constructed to evaluate the utility of the session as perceived by the students, for example whether they considered it a valuable learning opportunity and also whether the scenario was appropriate and felt realistic. Students were asked the degree to which they agreed with each statement using a 4-point rating scale ranging from 1 (not at all) to 4 (very much).

Procedure

Students attended the simulation individually at their allocated time-slot. They were first provided with an overview of the session content and given a tour of the simulation room to familiarise them with the lay-out and the SimMan 3G mannequin (the patient) [31]. Next, informed written consent was sought for inclusion in the research study and consenting students were allocated a unique ID number and randomly allocated to an experimental condition.

Those participating then filled in a 'pre-simulation questionnaire' which included the State and Trait portions of the STAI, the self-efficacy scale, and basic demographic information (age and sex). They were then fitted with the heart-rate monitor which was set up for each participant using their ID number, age, weight and height (controlled for at the data analysis stage).

The simulation then commenced. Students were introduced a senior doctor who provided them with a handover using a structured SBAR approach (Situation, Background, Assessment, Recommendation)

[32]. The experimental condition was manipulated as part of the handover (Fig 2). This same clinician observed and scored the participant's performance throughout the simulation through a two-way mirror.

Students entered the simulation room alone and began the ECG as they saw fit. Once all ECG leads had been attached the student was handed an ECG reading and told to use the phone to speak to a senior. The student's phone call was answered by the senior doctor who asked them what the ECG showed. Students were expected at their stage of training to identify 2 features of the trace: irregularity and tachycardia. If they did not mention these features or gave an incorrect interpretation they were coached towards the correct interpretation. Once irregularity and tachycardia had been confirmed the senior doctor asked the student to complete a drug calculation for preparation of a heparin drip (Fig 3) based on a potential diagnosis of atrial fibrillation. Students were asked to record their working on a sheet of paper and to return to the adjoining room once complete.

Immediately following the simulation participants removed their heart-rate monitor and completed a 'post-simulation questionnaire' including the State anxiety and self-efficacy items. They were then provided with a full debrief on their performance which included jointly establishing learning goals moving forward. Students were given the opportunity to ask questions and explore concerns. Following the debrief, a final 'post-debrief questionnaire' was completed including the State anxiety scale, self-efficacy scale, and session evaluation questions.

Data Analysis

All analyses were conducted using IBM SPSS Statistics 23. Descriptive statistics were computed for all variables. Histograms appeared normally distributed allowing parametric analysis. Mean heart-rate during the scenario was used as a summary measure for physiological stress. To assess the pattern of subjective stress overtime, total State anxiety scores were computed for each time-point (pre-sim, post-sim and post-debrief) and a repeated measures ANOVA was computed for the sample as a whole.

In line with the research questions, we then examined the relationship between physiological and psychological stress and performance for the sample as a whole. Correlations were conducted between stress measures (mean heart rate, and total state anxiety scores at each time-point) and the continuous performance measures (patient safety score, technical competence score, quality of communication score, total performance score, time spent talking and work rate). An independent samples t-test was computed to establish the difference in stress measures for those who did and did not engage in non-clinical conversation and perform and accurate drug calculation.

We then assessed the main effects and simple main effects of the experimental factors (urgency and cognitive load) for physiological and psychological stress, and all performance measures (as above), using one-way ANOVA (for continuous variables) and chi-square for binary variables.

Finally, we examined the relationship between trait anxiety and stress/performance, and self-efficacy and stress/performance using the same approach.

Session evaluation data were summarised descriptively.

RESULTS:

50 students volunteered to participate in the session (28% response rate). They had an average age of 20.6 years (sd = 0.4, range 18-33), and 39 (78%) were female.

Stress:

Psychological stress: State anxiety scores are summarised in Table 1. For all participants combined, there was a significant difference in State anxiety over time, whereby self-reported anxiety was significantly lower post-debrief than it was pre (mean difference 3.3, $p < 0.001$) or immediately post (mean difference = 3.8, $p < 0.001$) simulation. The same pattern of stress over time was observed when considering each experimental condition in isolation (F s range: 4.1-17.1, $p < 0.05$).

Physiological stress: For all participants combined the average mean heart rate during the scenario was 104.95 beats per minute (sd=18.63).

Table 1. Summary of State anxiety, Trait anxiety and self-efficacy scores

	All participants (n=50)	Low urgency/Low CL (n=13)	High urgency / Low CL (n=12)	Low urgency / High CL (n=12)	High urgency / High CL (n=13)
State Anxiety					
Pre-sim Mean (sd) Range	12.26 (2.9) 7-20	12.15 (2.3) 9-17	13.0 (2.7) 9-18	10.6 (2.5) 7-15	13.2 (3.6) 9-20
Post-sim Mean (sd) Range	13.14 (3.7) 6-23	13.15 (2.5) 10-17	14.25 (4.5) 7-23	11.8 (3.4) 7-20	13.3 (4.2) 6-20
Post-debrief Mean (sd) Range	9.02 (2.7)*** 6-18	10.64 (2.8)* 7-15	9.22 (1.6)** 7-12	7.5 (1.6)*** 6-10	8.9 (3.5)*** 6-18
Trait anxiety					
Trait anxiety Mean (sd)	39.6 (8.6)	37.8 (9.0)	39.9 (6.9)	38.9 (4.9)	41.5 (12.3)
Self-efficacy					
Pre-sim Mean (sd)	574.5* (113.2)	568.6 (84.2)	597.5 (83.5)	615.00 (129.6)	520.8
Post-sim Mean (sd)	521.4* (162.8)	516.9 (156.0)	546.3 (113.2)	587.5 (179.7)	441.9 (175.3)
Post-debrief Mean (sd)	616.07* (140.8)	590.5 (149.0)*	645.0 (78.4)*	698.00* (130.0)	556.9 (149.9)*

State anxiety scale: 6-24. Trait anxiety scale: 20-80. Self-efficacy scale: 0-900.

Within-group effects: ***= $p < 0.001$, **= $p < 0.01$, *= $p < 0.05$

Performance:

Descriptive statistics for performance measures are summarised for all participants combined in Fig 4 and Table 2. Compliance with standard patient safety behaviours was variable, for example with hand wash only being used by approximately half of the sample and a patient wrist band check being carried out by around three quarters. With regards to technical competence, 70% of the students placed all 10 ECG leads correctly, but correct interpretation of the ECG appeared more challenging, particularly regarding identifying tachycardia (completed successfully by only 32%, of the sample). 28% of students made an error in the drug calculation. Students were generally effective at communicating with the patient regarding the planned procedure (over 80% of the sample provided a clear explanation of how and why they were completing the ECG), but struggled more in communicating with a senior over the phone (44% of students did not introduce themselves and 26% did not use a structured approach to relay the ECG findings). Thirty six percent of the sample engaged in non-clinical conversation during the procedure. On average, the participants took just over 4 minutes to complete the ECG and just over 2 minutes to complete the drug calculation.

Table 2. Summary of scores for continuous performance measures

Measure	Mean (sd)
Work rate: Time taken to perform ECG in secs	253.06 (52.2)
Work rate: Time taken to perform drug calculation in secs	132.39 (66.5)
Percentage of time spent talking to the patient	12.1 (9.1)
Patient safety score (out of 4)*	3.04 (0.94)
Technical competence score (out of 14)*	11.86 (2.08)
Communication score (out of 6)*	4.7 (1.33)
Total performance score (out of 0-25) *	20.6 (2.9)

*Higher score=superior performance

Relationship between stress and performance:

Total technical competence was negatively correlated with State anxiety at all time-points: pre-simulation ($r=-.35$, $p=0.01$), post-simulation ($r=-.41$, $p=0.003$) and post-debrief ($r=-.54$, $p<0.001$), showing that those with higher psychological stress levels performed worse technically. Additionally, total performance score was negatively correlated with post-simulation and post-debrief State anxiety ($r=-.31$, $p=0.03$ and $r=-.45$, $p<0.001$ respectively). There was no relationship between physiological stress (mean heart rate) and any of the performance measures.

Impact of clinical urgency and cognitive load on stress and performance:

Stress: There was no effect of experimental condition on either psychological or physiological stress (F_s range: 0.01-2.7, $p>0.05$).

Performance: There was a main effect of cognitive load for percentage of time spent talking during the ECG – with those in the high cognitive load group speaking significantly more ($F=12.1$, $p<0.001$). Additionally, whether or not students spoke to the patient about non-clinical topics during the ECG was significantly associated with clinical urgency – those in the high urgency groups were more likely to speak about non-clinical topics than those in the low urgency groups ($X^2 = 5.56$, $p=0.018$). There were no additional main effects or simple main effects for clinical urgency or cognitive load for any of the performance measures.

Trait anxiety and self-efficacy:

Mean trait anxiety scores and total self-efficacy scores are summarised in Table 1.

Trait anxiety: Mean trait anxiety for the sample as a whole was 39.6 which is comparable to normative samples [33-34], and trait anxiety did not differ according to experimental condition. Trait anxiety was positively related to State anxiety at all time-points (r s range: .32-.62, p s<0.05). With regards to performance, there was a positive relationship between Trait anxiety and time spent completing the ECG ($r=0.4$, $p<0.01$). There were no other associations between trait anxiety and performance.

Self-efficacy: For the sample as a whole, total self-efficacy scores varied significantly over time ($F(2,78) = 21.9$, $p<0.001$) – they were lowest immediately post-simulation and highest post-debrief. The same effect over time was observed in each experimental condition considered in isolation.

Total self-efficacy was related to psychological stress. Pre-sim, post-sim and post-debrief self-efficacy scores were negatively correlated with State anxiety scores at all 3 time-points (r s range: -.29-.59, p s<0.05). There was no relationship between self-efficacy and physiological stress.

Pre-simulation self-efficacy was not related to any of the performance scores. However, those who got the drug calculation correct and those who completed it the quickest time had higher self-efficacy scores following the simulation (post-simulation: $F(1,49)=5.7$, $p<0.05$ and $r=-.41$, $p<0.01$ respectively, post-debrief: $F(1,49)=11.4$, $p<0.01$ and $r=-.35$, $p<0.05$ respectively).

Session evaluation

The session was received extremely favourably by the students in all respects (Table 3).

Table 3. Session Evaluation

Item	Mean (sd)
I feel I have made a positive contribution to my learning	3.9 (.28)
I would value more of these simulated learning opportunities	3.9 (.2)
I would recommend this session to others	3.9 (.24)
The debrief was useful for reflecting on my practice	3.9 (.3)
The task required of me was appropriate for my level of experience	3.7 (.48)
The scenario felt realistic	3.4 (.64)

Scale: 1 (not at all) – 4 (very much)

DISCUSSION

In this pilot study we examined the relationship between stress and clinical performance in 2nd year medical students undertaking a simulated ECG scenario. In an attempt to ‘unpick’ the relationship we manipulated two variables via experimental conditions: clinical urgency (a potential stressor) and cognitive load (which may mediate negative impacts of stress on performance). There was evidence of a behavioural response to the conditions. Those assigned to a high cognitive load group spoke significantly more during the scenario (in line with explaining what they were doing as they went along) and those in a high urgency group were significantly more likely to engage in social (non-clinical) conversation with their patient.- which we suggest was an attempt to put the patient at ease. We failed however to find any association between experimental condition and stress or other performance measures. Alongside a potential lack of power (due to a conservative sample size), we feel that this can be explained by the fact that *all* students experienced a moderate degree of stress and likely experienced high cognitive load due to the novelty of the task, which masked any potential effects of the experimental manipulations [24]. It was the first time they had experienced the simulated environment and had been observed and assessed individually on newly acquired clinical skills. Moreover, self-reported State anxiety levels were elevated for all students pre-simulation before any manipulation had taken place.

Whilst there was no effect of experimental condition, it was still possible to look at variation in stress levels in relation to performance for the group as the whole. In doing so we found that those students who reported greater stress levels (both before and after the simulation) performed worse technically – they made more technical errors and/or omissions than those who reported feeling less stressed. These findings support the notion that stress might impair performance, and are in line with previous research [5,9]. We can’t however attribute this to greater cognitive load in this study. Poorer communication skills and patient safety behaviours were related to higher self-reported stress levels following but not prior to the scenario, which makes it more difficult to attribute these performance deficits to the experience of stress itself. These findings have clear implications for patient safety and support investment in training programmes and workplace design interventions that are informed by human factors principles around stress management and the systems approach to understanding error.

We also measured trait anxiety and task self-efficacy to assess whether these factors were related to stress and performance and potentially mediated any impact of stress. We found no evidence of the latter. Both trait anxiety and task self-efficacy were related to State anxiety at each timepoint which was not unexpected given the strong relationship between these constructs reported in previous work [33,35-37]. Trait anxiety was also positively related to the amount of time taken to complete the ECG

(perhaps indicating that those with higher trait anxiety were more calculated in their approach) but it was not related to any other outcome variable. Task self-efficacy was lowest immediately post-simulation and highest immediately post-debrief for all students which highlights the importance of an effective debrief in addressing students' concerns (State anxiety was lowest post-debrief which also support this). Finally, task self-efficacy did not predict performance – scores prior to the simulation were not related to performance measures - however those who got the drug calculation correct and spent less time completing it reported higher self-efficacy post-simulation which was likely reflects their awareness of their success [38].

Whilst we revealed some important correlations between stress and performance here, the results of this study are also important from an educational perspective. Firstly, we revealed gaps in learning that require educational reinforcement. For instance, student compliance with some basic patient safety behaviours, which had recently been taught and assessed, was poor; interpretation of the ECG reading, even at a very basic level, proved challenging for many; and the use of a structured approach for communicating information to a senior was demonstrated by few. Secondly, until now, students at our medical school did not partake in simulation until the penultimate year (year 4) of their degree course. The current session was introduced as part of a curriculum re-design aimed at integrating stage-appropriate simulation opportunities earlier on. The positive feedback received and the ability of the students to engage appropriately suggests that this strategy was both feasible and beneficial.

Limitations, strengths and future directions

This study had certain limitations, which in part reflect the challenges of delivering robust research designs in a teaching environment. Firstly, physiological stress was measured using average heart-rate during the simulation. Whilst this gives an impression of arousal during the task, ideally we would have measured baseline heart-rate and calculated the increase, or physiological stress response for each individual in relation to their performance -however restrictions in the time available prevented us from doing so. Our measure of physiological stress was therefore fairly crude and likely lacked sensitivity in picking up variation between individuals. In future cohorts we plan to timestamp our measures to allow analysis of how heart-rate differs according to the specific task being performed (ECG vs drug calculation etc). Secondly, since attendance at the simulation was voluntary it is possible that the students who attended were not representative of the year and were skewed in terms of performance and/or anxiety, particularly given the modest response rate (28%). This will be addressed in future cohorts when attendance becomes mandatory. On a similar vein, one might question whether these findings were a virtue of involving medical students early on in their training and whether similar findings would be observed with more senior students or junior doctors. Having designed the project to be stage-appropriate i.e. assessing only those clinical skills that had already

been trained, we see no reason why the results should be particular to the group assessed. However, it would be interesting to apply the design to scenarios with more senior students going forward. Next, we undertook multiple statistical comparisons without imposing Bonferroni or similar corrections which increases the chance of Type 1 error. Whilst it is important to take this into account, we felt this was the appropriate approach to take at this exploratory/pilot stage to avoid missing potentially important effects which will inform hypothesis building moving forward. Finally, we feel that our manipulation of cognitive load may have been flawed. While previous experimental studies have typically increased cognitive load (burden on working memory) by requiring participants to engage in dual-processing tasks such as counting backwards, we wanted to ensure our manipulation held clinical relevance to preserve the fidelity of the scenario. We therefore asked students to articulate their clinical actions as they went along (an additional communication task). Whilst this likely increased cognitive load, 'thinking out loud' in this way also has the potential to enhance performance by allowing transition to a slower, more analytical decision-making style [39]. In future those in high cognitive load groups will be distracted by a series of questions by the patient which should increase demand on working memory but avoid thinking out loud.

Despite these limitations this study was one of few to endeavour to assess the relationship between stress and performance simultaneously using robust, real-time measures. Performance was observed rather than self-reported retrospectively; the outcomes were clinically relevant and included seldom assessed behaviours such as compliance with patient safety checks; validated measures were used where possible; and the simulated environment offered high clinical fidelity whilst also allowing a good degree of experimental control and standardisation.

Conclusion

In a population of medical students we have supported the emerging evidence-base suggesting that acute elevated stress can impair performance of clinical skills. This reinforces the importance of taking stress seriously in medical settings from a patient safety perspective, and putting systems in place to support effective stress management and awareness. Future work should continue to unpick which tasks are most vulnerable to interference from stress using high fidelity and well controlled methodologies, such as those afforded by simulation. From a medical educational perspective we would encourage those with a role in curriculum design to incorporate stage-appropriate introduction to simulation-based education right from the start of medical school.

Competing interests:

Dr S Russ has recently become a member of the Editorial board for BMJ Simulation and Technology Enhanced Learning. Aside from this we have no competing interests to declare.

Author Contributions:

All authors were involved in formulation of the research concept, research questions and study design. SR, IM, CB, JM & RM were involved in data collection. SR and RM undertook data analysis. SR & MJ drafted the manuscript and all authors were involved in editing it. All authors have approved the manuscript for submission.

Funding:

This study ran as part of the MBChB curriculum at the University of Aberdeen and no additional funding was required.

REFERENCES

1. Burbeck R, Coomber S, Robinson SM, Todd C. Occupational stress in consultants in accident and emergency medicine: a national survey of levels of stress at work. *Emergency Medicine Journal*. 2002 May 1;19(3):234-8.
2. Coomber S, Todd C, Park G, Baxter P, Firth-Cozens J, Shore S. Stress in UK intensive care unit doctors. *British journal of anaesthesia*. 2002 Dec 1;89(6):873-81.
3. Healy S, Tyrrell M. Stress in emergency departments: experiences of nurses and doctors: Sonya Healy and Mark Tyrrell review accounts of acute stress among healthcare professionals and show how it can be anticipated, reduced and managed. *Emergency nurse*. 2011 Jul 5;19(4):31-7.
4. Ramirez AJ, Graham J, Richards MA, Gregory WM, Cull A. Mental health of hospital consultants: the effects of stress and satisfaction at work. *The Lancet*. 1996 Mar 16;347(9003):724-8.
5. Kirkcaldy BD, Trimpop R, Cooper CL. Working hours, job stress, work satisfaction, and accident rates among medical practitioners and allied personnel. *International Journal of Stress Management*. 1997 Apr 1;4(2):79-87.
6. Sochos A, Bowers A, Kinman G. Work stressors, social support, and burnout in junior doctors: Exploring direct and indirect pathways. *journal of employment counseling*. 2012 Jun 1;49(2):62-73.
7. Heim E. Job stressors and coping in health professions. *Psychotherapy and psychosomatics*. 1991;55(2-4):90-9.
8. Thomas I, Nicol L, Regan L, Cleland J, Maliepaard D, Clark L, Walker K, Duncan J. Driven to distraction: a prospective controlled study of a simulated ward round experience to improve patient safety teaching for medical students. *BMJ quality & safety*. 2014 Nov 24;bmjqs-2014.
9. Houston DM, Allt SK. Psychological distress and error making among junior house officers. *British Journal of Health Psychology*. 1997 May 1;2(2):141-51.
10. Brennan N, Corrigan O, Allard J, Archer J, Barnes R, Bleakley A, Collett T, Bere D, Regan S. The transition from medical student to junior doctor: today's experiences of Tomorrow's Doctors. *Medical education*. 2010 May 1;44(5):449-58.
11. Cohen R A. Yerkes-Dodson Law. *Encyclopaedia of Clinical Neuropsychology* 2011. 2737-2738.
12. Qin S, Hermans EJ, van Marle HJ, Luo J, Fernández G. Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological psychiatry*. 2009 Jul 1;66(1):25-32.
13. McEwen BS, Sapolsky RM. Stress and cognitive function. *Current opinion in neurobiology*. 1995 Apr 30;5(2):205-16.
14. Crichton M, O'Connor P, Flin R. *Safety at the Sharp End: A Guide to Non-Technical Skills*. Ashgate Publishing, Ltd.; 2013 Aug 28.
15. Arora S, Sevdalis N, Aggarwal R, et al. Stress impairs psychomotor performance in novice laparoscopic surgeons. *Surgical Endoscopy* 2010; 24: 2588-93.
16. Ryan C, Ross S, Davey P, Duncan EM, Francis JJ, Fielding S, Johnston M, Ker J, Lee AJ, MacLeod MJ, Maxwell S. Prevalence and causes of prescribing errors: the PRescribing Outcomes for Trainee Doctors Engaged in Clinical Training (PROTECT) study. *PLoS One*. 2014 Jan 3;9(1):e79802.
17. Nielsen KJ, Pedersen AH, Rasmussen K, Pape L, Mikkelsen KL. Work-related stressors and occurrence of adverse events in an ED. *The American journal of emergency medicine*. 2013 Mar 31;31(3):504-8.
18. Flin R, O'Connor P, Crichton M. *Safety at the sharp end*. Cornwall, UK: Ashgate. 2008.
19. Angerer P, Weigl M. Physicians' psychosocial work conditions and quality of care: a literature review. *Professions and Professionalism*. 2015 May 13;5(1).
20. Ross S, Ryan C, Duncan EM, Francis JJ, Johnston M, Ker JS, Lee AJ, MacLeod MJ, Maxwell S, McKay G, McLay J. *BMJ quality & safety*. 2013 Feb 1;22(2):97-102.

21. Lewis PJ, Ashcroft DM, Dornan T, Taylor D, Wass V, Tully MP. Exploring the causes of junior doctors' prescribing mistakes: a qualitative study. *British journal of clinical pharmacology*. 2014 Aug 1;78(2):310-9.
22. Young JQ, Van Merriënboer J, Durning S, Ten Cate O. Cognitive load theory: Implications for medical education: AMEE guide no. 86. *Medical teacher*. 2014 May 1;36(5):371-84.
23. Staal MA. Stress, cognition, and human performance: A literature review and conceptual framework.
24. Fraser K, Ma I, Teteris E, Baxter H, Wright B, McLaughlin K. Emotion, cognitive load and learning outcomes during simulation training. *Medical education*. 2012 Nov 1;46(11):1055-62.
25. Sorg BA, Whitney P. The effect of trait anxiety and situational stress on working memory capacity. *Journal of Research in Personality*. 1992 Sep 30;26(3):235-41
26. Grau R, Salanova M, Peiro JM. Moderator effects of self-efficacy on occupational stress. *Psychology in Spain*. 2001;5(1):63-74.
27. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*. 1992 Sep 1;31(3):301-6.
28. Rule WR, Traver MD. Test-retest reliabilities of State-Trait Anxiety Inventory in a stressful social analogue situation. *Journal of Personality Assessment*. 1983 Jun 1;47(3):276-7.
29. Bieling PJ, Antony MM, Swinson RP. The State--Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behaviour research and therapy*. 1998 Aug 1;36(7):777-88.
30. Bandura A. Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents*. 2006;5(307-337).
31. SimMan 3G Mannekin. <http://www.laerdal.com/us/doc/85/SimMan-3G>
32. Haig KM, Sutton S, Whittington J. SBAR: a shared mental model for improving communication between clinicians. *The joint commission journal on quality and patient safety*. 2006 Mar 1;32(3):167-75.
33. Spielberger CD, Gorsuch RL, Lushene RE, Vagg PR, Jacobs GA. *Manual for the State—Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press, 1983
34. Kvaal K, Ulstein I, Nordhus IH, Engedal K. The Spielberger state-trait anxiety inventory (STAI): the state scale in detecting mental disorders in geriatric patients. *International journal of geriatric psychiatry*. 2005 Jul 1;20(7):629-34
35. Tang J, Gibson SJ. A psychophysical evaluation of the relationship between trait anxiety, pain perception, and induced state anxiety. *The Journal of Pain*. 2005 Sep 30;6(9):612-9.
36. Marquez DX, Jerome GJ, McAuley E, Snook EM, Canaklisova S. Self-efficacy manipulation and state anxiety responses to exercise in low active women. *Psychology and Health*. 2002 Jan 1;17(6):783-91.
37. Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall, Englewood Cliffs, NJ.
38. Bandura, A. (1988). Self-efficacy conception of anxiety. *Anxiety Research*, 1(2), 77–98.
39. Moulton CA, Regehr G, Lingard L, Merritt C, MacRae H. Slowing down to stay out of trouble in the operating room: remaining attentive in automaticity. *Academic Medicine*. 2010 Oct 1;85(10):1571-7.