

CLINICAL AUDIT

Dynamic confidence during simulated clinical tasks

A J Byrne, M T Blagrove, S J P McDougall

Postgrad Med J 2005;81:785–788. doi: 10.1136/pgmj.2004.029942

Objective: Doctors' confidence in their actions is important for clinical performance. While static confidence has been widely studied, no study has examined how confidence changes dynamically during clinical tasks.

Method: The confidence of novice (n = 10) and experienced (n = 10) trainee anaesthetists was measured during two simulated anaesthetic crises, bradycardia (easy task) and failure to ventilate (difficult task).

Results: As expected, confidence was high in the novice and experienced groups in the easy task. What was surprising, however, was that confidence during the difficult task decreased for both groups, despite appropriate performance.

Conclusions: Given that confidence affects performance, it is alarming that doctors who may be acting unsupervised should lose dynamic confidence so quickly. Training is needed to ensure that confidence does not decrease inappropriately during a correctly performed procedure. Whether time on task interacts with incorrect performance to produce further deficits in confidence should now be investigated.

The amount of confidence that a doctor feels has important implications for their willingness to conduct procedures, their willingness to ask for advice or a second opinion, and for their assessment of their skills.¹ The investigation of confidence is thus important, but studies differ in how confidence is measured and in their definitions of confidence, which range from confidence in one's general role² to confidences for performing specific tasks.³

Previous research has studied two main types of task specific confidence:

(1) Confidence in how one would perform certain procedures.

Marteau *et al*⁴ asked doctors to rate "how confident they felt when performing resuscitation." The number of cardiac arrests they had attended over the previous six months was found to be significantly correlated with level of confidence, but not with resuscitation skills as assessed during a training programme. Morgan and Cleave-Hogg⁵ assessed "confidence in ability to manage patient problems", these 25 problems being taken from various clinical experiences, and similarly Hutchinson and Robson⁶ used a five point rating of confidence (1 = low, 5 = high) in various areas of knowledge and skills, with mean scores for different procedures ranging from 2.6 to 4.0. This anticipatory type of confidence also includes confidence as "a judgement which influenced whether or not to undertake an activity", the definition used by Stewart *et al*.⁷ These confidence ratings can have, however, only low correlations with performance accuracy.⁵

(2) Confidence as self assessment during current performance.

Fitzgerald *et al*⁸ studied actual and self assessed examination performance over three years, finding accurate self

assessment of knowledge in years 1 and 2, but overconfidence in self assessment in year 3, when the examination was for clinical skills. Lewis *et al*⁹ found that more experienced doctors had higher confidence when correctly answering medical questions than did less experienced doctors, and that sleep loss led to increases in confidence about performance in answering medical questions, but a decrease in overall confident mood.

We wished to advance the work on task specific confidence by investigating changes in task specific confidence during clinical situations, in contrast with the simpler task specific confidence assessments mentioned above that necessitate only a single response for the whole task. We thus needed to obtain a sequence of responses from the trainees, and to ask for confidence ratings at intervals during the procedure.

For this we used two crisis scenarios on an anaesthetic simulator that each last some minutes and where both the clinical problem and the solutions should be familiar to the trainees, who were thus likely to be making correct responses throughout. Simulation was chosen as it has been found to be highly rated by students for the learning experience, for its appropriate content, and for its use as an evaluation tool.⁵ The tasks were bradycardia and failure to ventilate, these being classed as easy and as difficult respectively by a survey of consultant anaesthetists' perception of a variety of clinical problems.

METHOD

The study was given ethical approval by the Swansea NHS Trust Ethical Approval Committee. Twenty trainee anaesthetists (male = 12, female = 8) were recruited from a single hospital and gave informed consent to take part. The subjects completed two short simulations containing a clinical problem using an intermediate fidelity ACCESS simulator.^{10 11} Half the trainees were novices with less than 12 months' experience and were working with close clinical supervision. The "experienced" group were also trainees, but who had more than 12 months' experience and were working in a post without direct supervision.

The clinical scenario for both simulations involved an 80 year old, diet controlled diabetic "patient" undergoing an internal reduction and fixation of a complex ankle fracture. The patient had been anaesthetised, paralysed, and intubated and was in the middle of a prolonged procedure. The first simulation was intended to reproduce a simple problem that would respond to a simple solution that should have been familiar to all the trainees. It was therefore complicated by a simple bradycardia that responded rapidly to a single dose of intravenous anticholinergic drug. The second simulation was intended to reproduce a severe problem that would not respond to any simple measures. The scenario chosen was a sudden inability to ventilate the patient, associated with a progressive fall in oxygen saturation, in an otherwise stable patient. The clinical signs were considered to be compatible with four possible diagnoses—physical blockage of the airway, failure of the ventilator, severe bronchospasm, and bilateral tension pneumothorax. In keeping with a diagnosis

of severe refractory bronchospasm, the high resistance to ventilation was continued for five minutes and did not respond to any of the trainees' actions. To avoid any adverse psychological effect on the trainees, the airway resistance was reduced to normal after five minutes, as if the problem had resolved. It was expected that the trainees would interpret any improvement as a response to their last intervention.

Confidence levels were measured by asking each trainee to rate verbally their own confidence on a five point scale every minute during the simulation. Each trainee was given guidance on the rating of confidence before the simulations. They were told to report a confidence level of five when they felt that they understood the condition of the patient and felt able to cope with these problems. A confidence level of one showed a complete lack of understanding of the situation and inability to deal with the situation. They were also advised not to consider their level of anxiety, so that even if they felt anxious this should not be interpreted as a lower level of confidence. The use of interventions during simulation to measure performance has been used previously in simulation and is not thought to significantly change the performance of subjects.¹²

Each simulation began with a single minute period of normality and ended once the trainees' confidence levels had returned to normality, or two minutes had passed since the clinical problem had resolved. The operator during each simulation recorded a written note of events. All the simulations were recorded by a single camcorder and stored as a digital recording for later analysis, as used previously.¹³ The actions of each trainee were noted during the simulation and checked later with the video recordings. As with any study of this type, it is impossible to accurately determine the reasons for each of the trainees' actions, as we are not able to determine their perception of the problem at the time the action was performed. It was therefore assumed that the actions of the trainees were appropriate to their provisional diagnosis and classified into three types, correct, non-specific, and wrong. Correct actions were those that could have led to the solution of the presumed problem (shown in parentheses below). In the first simulation the "correct" action was the administration of an anticholinergic drug (bradycardia). In the second scenario, "correct" actions were considered to be removing the tracheal tube (blocked tube), manual ventilation (ventilator failure), changing to an alternative method of ventilation (blockage of the circuit), needle aspiration of the chest/percussion (identify/treat tension pneumothorax), and administration of a bronchodilator (bronchospasm). Non-specific interventions were those interpreted as not treating the immediate problem, but not detrimental to the "patient",

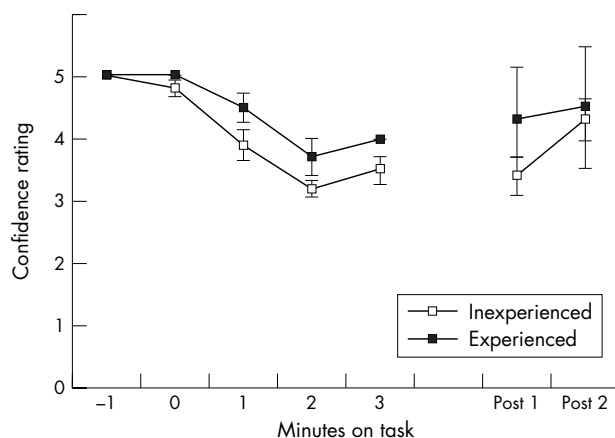


Figure 1 Confidence levels during the easy task (bradycardia).

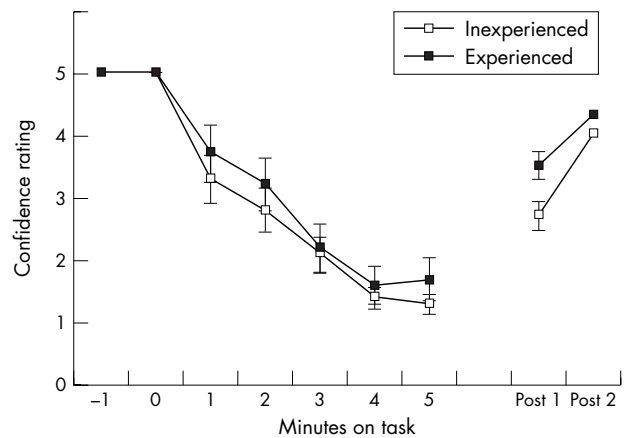


Figure 2 Confidence levels during the difficult task (failure to ventilate).

for example, starting intravenous fluid infusion or increasing the inspired oxygen concentration. "Wrong" actions were those actions that might have been detrimental to a real patient, for example, giving an inappropriate drug.

Confidence levels in each group were expressed as the median confidence at each time point for the experienced and novice groups.

The number of correct, non-specific, and wrong treatments given by each trainee was noted as well as the number of presumed specific diagnoses. A single rater was used as Morgan and Cleave-Hogg¹⁴ found a very high inter-rater reliability for independent assessment of videoed performance ($r = 0.87$).

RESULTS

All the simulations were completed successfully with data available in all cases. In two cases, one from each group, the video recording equipment failed to record events. This may have resulted in some of the actions of two trainees not being recorded.

In the first simulation (bradycardia) all the subjects treated the "patient" with intravenous atropine as expected. Non-specific actions noted were the administration of an inotrope (intravenous ephedrine/methoxamine) or intravenous fluids, increasing the inspired oxygen concentration, reducing the inspired anaesthetic concentration, and switching to manual ventilation. The novice group completed a median (range) of 3 (2-5) non-specific actions while the experienced group completed a median (range) of 3 (0-5). Confidence during the first simulation remained high with the experienced group showing higher confidence levels (see fig 1).

In the second simulation all the subjects recognised the problem and started appropriate treatments. The novice group performed a median (range) of 3 (2-4) actions, with all those tested choosing to switch to manual ventilation and to remove the tracheal tube. Five trainees administered a bronchodilator and six used a self inflating bag as an alternative means of ventilation. The experienced group performed a median (range) of 4 (3-5) actions. Almost all the experienced group chose to manually ventilate the patient, administer a bronchodilator, remove the tracheal tube, and to use a self inflating bag, but only four percussed the chest or aspirated the chest with a needle. Confidence levels were lower during the second simulation with little difference between the two groups (see fig 2).

An analysis of variance was carried out to examine the effects of level of experience (novice compared with experienced), type of task (easy compared with difficult), and the time at which confidence ratings were obtained. As

the time taken to carry out the two tasks differed, average confidence ratings were obtained for the beginning of the task (at -1 and 0 minutes), the middle of the task (1-3 minutes for the easy task and 1-5 minutes for the hard task), and the end of the task (1 and 2 minutes after task). The effects of time were therefore considered at three levels (beginning compared with middle compared with end). Overall, the experienced group were more confident than the novice group, (M (experienced) = 4.15, M (novice) = 3.81; $F(1,18) = 4.65$, $MSE = 0.73$, $p < 0.05$) and confidence ratings were higher for the easy than for the difficult task (M (easy) = 4.31, M (difficult) = 3.65; $F(1,18) = 110.43$, $MSE = 0.12$, $p < 0.001$). Ratings also changed over time with a decrease in ratings apparent in the middle of the task (M (beginning) = 4.97, M (middle) = 3.09, M (end) = 3.87; $F(1,18) = 110.43$, $MSE = 0.12$, $p < 0.001$).

Interestingly, there was an interaction between task and time ($F(2,18) = 36.44$, $MSE = 0.18$, $p < 0.05$; see fig 2). Analyses of simple main effects showed that ratings were equally high at the beginning of both tasks ($F(1,18) = 2.25$, $p > 0.05$) but confidence ratings were higher for the easy task compared with the difficult task in the middle ($F(1,18) = 138.08$, $p < 0.001$) and at the end ($F(1,18) = 8.57$, $p < 0.01$). It is important to note, however, that the effect size was large during the middle of the task ($\eta^2 = 0.88$) and only moderate by the end ($\eta^2 = 0.33$) suggesting that differences in ratings were gradually returning to uniformly high levels. No other interactions were significant.

Finally, separate analyses of variance were carried out to examine the magnitude of the task by time interaction for the experienced and novice groups. When trainees were experienced differences in ratings were more apparent between tasks ($F(2,18) = 70.27$, $MSE = 0.30$, $p < 0.001$) and the effect size was large ($\eta^2 = 0.88$). For novice trainees, differences in ratings between tasks were less and the effect size was moderate ($F(2,18) = 9.65$, $MSE = 0.05$, $p < 0.01$, $\eta^2 = 0.52$). This suggests that differences between tasks were more pronounced for the experienced than the novice group.

DISCUSSION

Confidence is important to clinicians, as the complexity of the clinical environment requires them to rapidly diagnose problems and institute treatments. This study shows that it is possible to measure the confidence levels of anaesthetic trainees during simulated clinical practice and that their level of confidence is related both to their experience and inversely to the severity of the problem.

The results of the novice trainees are unsurprising as they were able to recognise and treat the bradycardia with considerable confidence, but rapidly lost confidence when faced with the more difficult problem. The more experienced trainees were more confident when faced with the bradycardia, which could easily be predicted.

However, the responses of the experienced trainees to the failure to ventilate was surprising as they showed a considerable loss of confidence within four minutes of the onset of the problem that was not significantly different from the novice trainees. This accords with the finding of Whitehouse *et al.*,² who, having assessed confidence at the end of a training course, state that "confidence may evaporate when faced with real experience".

Bronchospasm is not an uncommon clinical problem and it was surprising that a group of comparatively experienced clinicians should lose confidence so quickly. Our interpretation of these results is that in a complex clinical situation, clinicians rely heavily on pattern recognition. When a problem occurs, a pattern is recognised and a "stock" solution is used. This view is supported by the finding that the experienced group took fewer actions in response to the

bradycardia (increased diagnostic accuracy) and a larger number of actions in response to the difficult task (wider range of options available). Experienced staff are able to recognise problems and institute a wider range of solutions with little stress, which reinforces their high level of confidence. However, when faced with a problem that did not immediately respond, confidence levels fell.

We propose that it is the uncertainty of what will happen next that is causing the lack of confidence in the trainees. Hall¹⁵ reviews how uncertainty can be denied during much medical decision making, and reviews responses to uncertainty that impose apparent clarity on the situation. In contrast, we have found the opposite reaction to uncertainty in our trainees, a steady decrease in confidence despite methodical and correct actions.

Appropriately low confidence can be useful, for example, Hays *et al.*¹ review how the training of insight into one's own performance can provoke improvement in poor performance. The results of this study show, however, that confidence can steadily decrease in trainee anaesthetists in a situation where their performance is appropriate, simply because the problem has not yet been solved. This is important as lack of confidence may itself adversely affect performance,¹⁶ leading to further loss of confidence, worsening performance, and defensive practice.¹⁷

Having found this effect of time on task on confidence, a further study is now needed in which some responses made are wrong and some are right: at issue is whether confidence in incorrect decisions will follow the same path as the confidence in correct decisions that was investigated here. Morgan and Cleave-Hogg³ conclude that it is important to develop strategies for students to accurately judge their capabilities to undertake various clinical procedures, because such level of confidence would then predict performance. Our results show that this assumption may be too simple, in that confidence can change during a procedure.

Clearly, a training intervention is needed to ensure that students' technical ability and confidence rise progressively and in concert. It seems probable that appropriate dynamic confidence cannot be gained during unsupervised clinical practice and can only be gained by allowing students to practise their skills within a controlled environment, with both technical support and the opportunity to objectively review and reflect on their own performance. This training could either be provided by using simulations within an educational environment, or by permitting supervised practice in the clinical environment. The ability to review performance against established criteria with the aid of a supervisor has already been used to overcome overconfident "impression management" by Evans *et al.*,¹⁸ and change toward the use of simulation and supervised clinical practice has already been recommended.¹⁹ Our suggestion is that dynamic confidence should be assessed during such tasks. Students with low dynamic confidence can then be reassured with active support, and instructed to focus on whether their current actions are appropriate, rather than focusing on possible future problems. Also, as people with higher non-work related psychological distress have lower confidence scores overall and smaller increases in confidence across training sessions,²⁰ sources and effects of current non-work related distress may need to be investigated. In contrast, those with high levels of confidence can be challenged with more difficult material. The aim would be to train, not just until technical success, but to the point where the trainee also feels confident of success. This training would bear in mind that confidences for particular procedures may differ from each other and have different rates of change across the years of training,²¹ and that inappropriate confidence is more common in difficult than in easier tasks.²²

Authors' affiliations

A J Byrne, Swansea Clinical School, Morriston Hospital, Swansea, Wales

M T Blagrove, S J P McDougall, Department of Psychology, University of Wales Swansea, Swansea, Wales

Funding: none.

Competing interests: none declared.

Correspondence to: Dr A J Byrne, Swansea Clinical School, Morriston Hospital, Swansea SA6 6NL, Wales; aidan.byrne@ntlworld.com

Submitted 21 October 2004

Accepted 15 March 2005

REFERENCES

- 1 **Hays RB**, Jolly BC, Caldon LJM, *et al*. Is insight important? Measuring capacity to change performance. *Med Educ* 2002;**36**:965–71.
- 2 **Whitehouse CR**, O'Neill P, Dornan T. Building confidence for work as house officers: student experience in the final year of a new problem-based curriculum. *Med Educ* 2002;**36**:718–27.
- 3 **Owen H**, Plummer JL. Improving learning of a clinical skill: the first year's experience of teaching endotracheal intubation in a clinical simulation facility. *Med Educ* 2002;**36**:635–42.
- 4 **Marteau TM**, Wynne G, Kaye W, *et al*. Resuscitation: experience without feedback increases confidence but not skill. *BMJ* 1990;**300**:849–50.
- 5 **Morgan PJ**, Cleave-Hogg D. Comparison between medical students' experience, confidence and competence. *Med Educ* 2002;**36**:534–9.
- 6 **Hutchinson S**, Robson WP. Confidence levels of PRHOs in caring for acutely ill patients. *Postgrad Med J* 2002;**78**:697.
- 7 **Stewart J**, O'Halloran C, Barton JR, *et al*. Clarifying the concepts of confidence and competence to produce appropriate self-evaluation measurement scales. *Med Educ* 2000;**34**:903–9.
- 8 **Fitzgerald JT**, White CB, Gruppen LD. A longitudinal study of self-assessment accuracy. *Med Educ* 2003;**37**:645–9.
- 9 **Lewis KE**, Blagrove M, Ebdon P. Sleep deprivation and junior doctors' performance and confidence. *Postgrad Med J* 2002;**78**:85–7.
- 10 **Byrne AJ**, Hilton P, Lunn J. Basic simulations for anaesthetists: a pilot study of the ACCESS system. *Br J Anaesth* 1994;**49**:376–81.
- 11 **Byrne AJ**, Jones JG. Responses to simulated anaesthetic emergencies by anaesthetists with different durations of clinical experience. *Br J Anaesth* 1997;**78**:553–6.
- 12 **Weinger MB**, Herndon OW, Gaba DM. The effect of electronic record keeping and transesophageal echocardiography on task distribution, workload and vigilance during cardiac anesthesia. *Anesthesiology* 1990;**87**:144–55.
- 13 **Byrne AJ**, Jones JG. Inaccurate reporting of simulated critical anaesthetic incidents. *Br J Anaesth* 1997;**78**:637–41.
- 14 **Morgan PJ**, Cleave-Hogg D. Evaluation of medical students' performance using the anaesthesia simulator. *Med Educ* 2000;**34**:42–5.
- 15 **Hall KH**. Reviewing intuitive decision-making and uncertainty: the implications for medical education. *Med Educ* 2002;**36**:216–24.
- 16 **Trumbo D**, Noble M. Response uncertainty in dual-task performance. *Organ Behav Hum Perform* 1972;**7**:203–15.
- 17 **Passmore K**, Leung W-C. Defensive practice among psychiatrists: a questionnaire survey. *Postgrad Med J* 2002;**78**:671–3.
- 18 **Evans AW**, Leeson RMA, Newton-John TRC. The influence of self-deception and impression management on surgeons' self-assessment scores. *Med Educ* 2002;**36**:1095.
- 19 **General Medical Council**. *Tomorrow's doctors*. London: GMC, 2002.
- 20 **Williams S**, Dale J, Glucksman E, *et al*. Senior house officers' work related stressors, psychological distress, and confidence in performing clinical tasks in accident and emergency: a questionnaire study. *BMJ* 1997;**314**:713–18.
- 21 **Marel GM**, Lyon PM, Field MJ, *et al*. Clinical skills in early postgraduate medical trainees: patterns of acquisition of confidence and experience among junior doctors in a university teaching hospital. *Med Educ* 2000;**34**:1013–15.
- 22 **Kebell MR**, Wagstaff GF, Covey JA. The influence of item difficulty on the relationship between eyewitness confidence and accuracy. *Br J Psychol* 1996;**87**:653–62.

bmjupdates+

bmjupdates+ is a unique and free alerting service, designed to keep you up to date with the medical literature that is truly important to your practice. bmjupdates+ will alert you to important new research and will provide you with the best new evidence concerning important advances in health care, tailored to your medical interests and time demands.

Where does the information come from?

bmjupdates+ applies an expert critical appraisal filter to over 100 top medical journals. A panel of over 2000 physicians find the few 'must read' studies for each area of clinical interest.

Sign up to receive your tailored email alerts, searching access and more...

www.bmjupdates.com