McRobert, A, Causer, J, Vassiliadis, J, Watterson, L, Kwan, J and Williams, MA

Contextual information influences diagnosis accuracy and decision making in simulated emergency medicine emergencies

Contextual information influences diagnosis accuracy and decision making in simulated Emergency Medicine emergencies

Allistair P. McRobert & Joe Causer

Liverpool John Moores University

John Vassiliadis & Leonie Watterson

Royal North Shore Hospital, Sydney

James Kwan

Westmead, Sydney

A. Mark Williams

Brunel University, London

Please address correspondence to:
Allistair McRobert
Liverpool John Moores University
Tom Reilly Building
Byrom Street
Liverpool, L3 3AF
United Kingdom
Tel - +44 151 904 6258
Email: a.p.mcrobert@ljmu.ac.uk

Keywords: Decision-making; emergency department; simulation
Word count (excluding title page, abstract, references, figures and tables): 3825

Contextual information influences diagnosis accuracy and decision making in simulated Emergency Medicine emergencies

RUNNING HEADER: Quality and context in decision-making

ABSTRACT

Background: It is well documented that adaptations in cognitive processes with increasing skill level support decision-making in multiple domains. We examined skill-based differences in cognitive processes in emergency medicine physicians and whether performance was significantly influenced by the removal of contextual information related to a patient's medical history.

Method: Skilled (*n* = 9) and less skilled (*n* =9) emergency medicine physicians responded to high fidelity simulated scenarios under high and low context information conditions.

Results: Skilled physicians demonstrated higher diagnostic accuracy irrespective of condition and where less affected by the removal of context-specific information compared to less skilled physicians. The skilled physicians generated more options, and selected better quality options during diagnostic reasoning compared with less skilled counterparts. These cognitive processes were active irrespective of the level of context–specific information presented, although high context information enhanced understanding of the patients' symptoms resulting in higher diagnostic accuracy.

Conclusion: Our findings have implications for scenario design and the manipulation of contextual information during simulation training.

INTRODUCTION

Medical errors are associated with significant patient morbidity and mortality worldwide [1]. The causes of these errors are multifactorial and include physician, patient, institutional and environmental factors. As in other fields of practice, there are individual differences in medical professionals' abilities, with some never attaining an expert level [2]. Emergency physicians engage in time-pressured judgements about critically ill patients frequently with limited information. Such complex decision-making tasks require emergency physicians to observe, recognise, encode and analyse a wide range of information before formulating an appropriate response to the clinical problem. These processes are performed in an environment where there are multiple and simultaneous distractions and where critical decisions are made hundreds of times in a day, with low tolerance for errors.

Skilled performers develop strategies for identifying relevant visual and contextual information sources that facilitate improved decision-making and performance [3]. Skilled law enforcement officers generate more relevant thought processes that positively correlate with decision quality/accuracy compared to their less skilled counterparts [4]. Similarly, Cormier *et al*. [5] examined the thought processes and performance of low and high performing student nurses during high fidelity medical simulations. They measured verbal statements coded as 'observations' (e.g., context, patient and monitor) or 'actions' (e.g., based on orders or patient condition) and physiological data of the patient's respiratory status. High performing student nurses are better at observing relevant cues suggesting patient deterioration and forward plan response actions that positively alter the physiological trajectory of the patient. Low performing student nurses verbalise more irrelevant information cues and fail to perform response actions directly related to the patient's condition.

Decision-making is further constrained by previous exposure to the environment and the contextual information provided by the situation. Perceptual-cognitive process (i.e., visual search strategies and thought processes) and performance (i.e., response accuracy and/or response time) measures in sport alter when contextual information about the environment and opponent is available compared to when this information is not present [6-8]. In the current study, we examine how contextual information influences performance and cognitive processing during simulated medical emergency scenarios.

Ericsson and Kintsch [9] proposed long-term working memory (LTWM) theory to account for how skilled individuals encode and retrieve relevant information rapidly from long-term memory (LTM). As a result of accumulated deliberate practice, key features or patterns from previous events are established as retrieval cues that, if activated, will rapidly retrieve large portions of complex information stored in LTM [see 10]. This information remains active during the task so that the individual can rapidly update their understanding of the current situation with contextual information that becomes available [11]. Furthermore, an individual uses this information to predict future events, evaluate multiple response options and rank how 'likely they are to succeed', consequently enhancing decision quality [12]. The collection of think-aloud verbal report protocols has demonstrated that the superior performance of skilled individuals is associated with the generation of more evaluation, prediction and planning statements compared to less skilled individuals [4,7,13,14].

The processing of contextual information is thought to be an essential component of decision-making because it activates retrieval structures and associated knowledge in LTM that relate to the situational demands of the task [9]. Verkoeijen *et al*. [15] demonstrated that expert physicians processed laboratory data faster and increased diagnosis accuracy when they were embedded in the clinical context (e.g., patient medical history and results of the physical examination). An increase in accuracy and reduction in response time have been reported in athletes when exposed to test conditions comparable to the sporting arena [6,16].

McRobert *et al*. [7] manipulated contextual information during a simulated cricket batting task. Participants were instructed to anticipate ball destination when viewing life-size video clips during two conditions. In a low-context condition they viewed randomised clips from different bowlers. In a high-context condition they viewed a sequence/block of six trials from the same opponent that provided cumulative information on the opponent's performance tendencies and replicated the competition situation. Skilled participants demonstrated superior anticipation performance, a more effective visual search strategy and verbalised more evaluation, prediction and deep planning statements across both conditions. All participants improved anticipation accuracy, made more evaluations and planned more future options when contextual information was made available. These findings support previous research demonstrating that decision-making improves with a practitioner's level of baseline experience and skill. Furthermore, these authors suggest that irrespective of participant skill level decision-making performance is enhanced when context rich information is included in the test conditions.

These findings are relevant to performance testing in the healthcare domain. In particular, emerging simulation technologies enable more manipulation of contextual information than previously afforded by traditional skills laboratories and testing methods. High fidelity patient simulators allow experimental control and replication of scenarios, while maintaining ecological validity and are widely accepted in the medical domain for assessing clinical practice, facilitating learning and discriminating performance differences between skilled and less skilled medics and health practitioners [17-19]. In this study, we employ a high fidelity simulation environment to measure performance (i.e., diagnosis accuracy), option generation strategies, and cognitive processes using concurrent verbal reports. We use a novel approach for manipulating context. In a high-context condition, participants have access to all information and resources, whereas in a low-context condition pre-determined information and resources are delayed (until after the scenario terminated) or removed during the experimental condition (i.e., low-context).

We hypothesised that skilled emergency medicine doctors would outperform their less skilled counterparts on diagnosis accuracy [20]. We hypothesised, based on LTWM theory [9], that skilled participants' superior diagnosis accuracy would be supported by more evaluation, prediction, and deep planning statements. In addition, denying vital contextual information related to the scenario would not allow participants to update their knowledge of the scenario. Therefore, we predict that skilled participants' decrement in diagnosis accuracy would result in fewer evaluation, prediction and deep planning statements in the low-context condition. According to LTWM theory, the skilled participants should generate more diagnosis options and select the best option (i.e., decision quality) later in the sequence, compared to less skilled participants.

METHOD

Participants

Participants were 9 skilled (mean age = 45.8 years, *SD* = 8) and 9 less skilled (mean age = 33.1 years, *SD* = 7.6) emergency medicine doctors. Skilled participants were Emergency Physicians with a mean of 20.1 years (*SD* = 9.3) experience. Less skilled participants were at a Registrar level (Emergency Medicine trainees) with a mean of 7.6 years (*SD* = 3.2) experience. The Northern Sydney Central Coast Area Health Service Human Research Ethics Committee approved the study.

Simulated Task Environment

The simulated task environment consisted of a Laerdal SimMan human patient simulator (Laerdal, Stavenger, Norway) and a simulated Emergency Department that contained many features of an actual Emergency Department including a bed, headboard and controls, realistic functioning medical gases, intravenous (IV) pumps, a real resuscitation cart, medications and medical supplies. Monitor displays reproduced vital signs measured in the Emergency Department. A console operator adjusted clinical signs based on scripts detailing scenario progress together with the simulator's real-time response to treatments. The scenarios were overseen from a soundproof observation room behind a two-way mirror by the clinicians involved in the study, each having in excess of 10 years experience as simulation practitioners. These individuals controlled the realism of the evolving scenario using previously agreed scripts and their clinical knowledge. Furthermore, the simulated task environment contained 2-angle video cameras and lapel microphones for capturing video and audio during scenarios.

Scenarios

Two emergency medicine scenarios were developed based on actual cases admitted to an emergency department. Both scenarios began with the simulated patient being handed over to either an Emergency Physician or a Registrar by a Simulation Fellow who was playing the role of a Registrar. The handover was scripted as was the request for help and did not vary from participant to participant. The Simulation Fellow would provide information only as allowed by the script. Each scenario started from the period the emergency doctor entered the simulated task environment and terminated after 20 minutes had elapsed. During both scenarios, a nurse was in the simulated Emergency Department providing information and assisting the participants as required. The nurse's responses were scripted and identical for all participants. Both faculty members in the room participated as team-member and were equipped with a two-way radio transceiver so that experimenters could standardise information provided to the participant involved in the scenario (e.g., pathology results from the laboratory).

In scenario one, the patient was a 19-year-old male who self presented with abdominal pain and vomiting. The underlying cause requiring diagnosis is rupture of previously ingested heroin-filled balloons on a return trip from Thailand. The participant is asked by the Registrar to assist with the assessment and management of this patient. The case evolves with the simulated patient becoming unconscious and stops breathing; a life threatening situation requiring immediate resuscitation. The history of a recent trip to Thailand and the abdominal x-ray results are provided as additional contextual information.

In the second scenario, a 34-year-old female who is 24-weeks pregnant self presents with right upper quadrant (RUQ) pain and severe shortness of breath (SOB) that has been on-going for a few hours. She has been diagnosed with a clotting abnormality and non-viable pregnancy for which a termination has been scheduled and the blood thinning medication ceased five days earlier. The Simulation Fellow hands her over and asks the participant for help to assess and manage her. The simulated patient is unwell and requires immediate intervention to manage a rapid heart beat, breathing difficulties and abdominal pain which is not responding to usual doses of narcotics. Investigation include chest x-ray and ultra-sound. A computed tomography (CT) and blood tests that provide further relevant information on liver infarcts and a typical patter consistent with the obstetric condition haemolytic anaemia, elevated liver enzymes and low platelet count (HELLP) syndrome are available as contextual information.

Procedure

Participants completed an information sheet regarding their biographical background and provided informed consent. Ericsson and Kirk's [21] adapted direction for giving think-aloud verbal reports required participants to provide verbal reports during a warm-up task containing non-clinical problems [22]. Participants practiced giving verbal reports with feedback for approximately 30 minutes prior to the medical scenarios. They were familiarised with the simulated task environment, instructed to provide concurrent verbal reports when not communicating with the patient or staff and given the opportunity to ask questions and interact with the simulator prior to the scenarios. Concurrent verbal reports were selected as they provide a more complete cognitive representation of performance within a simulated task environment, compared to retrospective verbal reports [23].

Prior to the experimental scenarios, lapel microphone and radio transmitters were fitted to the participant. Participants completed both scenarios in either a low- or high-context condition. The scenarios were counterbalanced for skill level and viewing order to reduce order and learning effects. During the low-context condition, pre-selected, contextual-specific information was not provided or was delayed so that it would not arrive before the termination of the scenario. In scenario one, the abdominal x-ray and parent were not available. In scenario two, the CT scan was not available and the bloods were delayed. All specified information detailed in the scenario section was available to the participant in the high-context condition.

On entering the simulated task environment, participants received a current patient history from the Simulation Fellow that contained a standardised script. Participants interacted with the patient and staff; however, the experimenters controlled the information received depending on the context condition. If the participant remained silent for an extended period of time, the nurse was instructed to remind them to concurrently verbalise their thoughts. On termination of the scenarios, participants were asked to record their primary diagnosis and rank two differential diagnoses in written format. In addition, they were instructed to verbally rank their diagnosis and provide justifications for their decisions.

Data Analysis

Diagnosis Data

Diagnosis accuracy was defined as the number of correct primary diagnosis on each scenario. These data were recorded as a frequency count and reported as ratio and percentage scores to compare across groups and context conditions.

Verbal Report data

The scenario videos were captured and analysed using Studiocode, version 3.5 (Sportstec Ltd, Australia). This system is an analytical tool that allows the researcher to mark and code segments of video and audio into categories. Using the procedures outlined in Ericsson and Simon, [24] verbal reports were initially segmented using natural speech and syntactical markers.

Verbal statements were coded based on categories adapted from Ericsson and Simon's [24] original structure that was further developed by McRobert *et al*. [7]. Concurrent verbal statements were assigned to monitor, evaluation, prediction or deep planning categories (Table 1). Fifty verbal statements were selected at random to establish objectivity and reliability using the inter- and intra-observer percentage agreement formulas [25]. The primary experimenter coded the 50 statements on four separate occasions (i.e., prior to coding, after 6, 12, and 18 participants) with an intra-observer agreement range from 94% to 98%. Inter-observer agreement was assessed prior to coding (90%) the statements, and after 12 participants (93%) by an independent experimenter.

[Insert Table 1 here]

Three-way analysis of variance (ANOVA) was conducted with group (less skilled/skilled) as the between participant factor and context (low/high) condition and verbal statement type (monitor/evaluation/prediction/deep planning) as the within-participant factors. Partial eta squared (η_p^2) values are provided as a measure of effect size. When making comparisons between two means, Cohen's *d* measures are reported. Posthoc Bonferroni corrected pairwise comparisons are reported as follow-ups where appropriate.

Option-generation Data

Participant prediction statements were analysed separately to determine the option generation process. A prediction statement was included if it made reference to a potential diagnosis. Two-way ANOVAs with group as the between participant factor and context condition as the within participant factor were used to measure the total number of options generated and initial serial position of the option selected as the final diagnosis option. Initial serial position of the option was calculated as a ratio of the number of options generated prior to the selection of the primary diagnosis. A ratio value closer to zero indicates that the participant's primary diagnosis occurred earlier in the option generation process. In contrast, a value closer to one indicates that the primary diagnosis was generated later in the option generation process.

RESULTS

Diagnosis Data

Skilled participants reported the correct diagnosis in 14/18 (78% accuracy) scenarios, whereas the less skilled participants reported the correct diagnosis on 7/18 (39% accuracy) scenarios. In addition, both groups reported 14/18 correct diagnoses during the high- (78%) compared 7/18 during the low-context condition (39%). Finally, skilled participants reported 9/9 (100%) correct diagnoses in the high-context condition, compared to 5/9 (56%) correct diagnoses in the low-context condition. In contrast, less skilled participants reported 2/9 (22%) correct diagnoses in the low- compared to 5/9 (56%) correct diagnoses in the high-context condition.

Verbal Report Data

There were no significant Group × Context Condition × Verbal Statement Type ($F_{3,48}$ = .26, p = .85, η_p^2 = .02), Context Condition × Verbal Statement Type ($F_{3,48}$ = 1.99, p = .13, η_p^2 = .11), or Group × Context Condition interactions ($F_{1,16}$ = .67, p = .43, η_p^2 = .04). There was a significant Group × Verbal Statement Type interaction ($F_{3,48}$ = 3.06, p = .04, η_p^2 = .16). Skilled participants made more evaluation (*d* = 1.01), prediction (*d* = 1.83) and deep planning (*d* = .70) statements than less-skilled participants, whereas there was no difference between the number of monitoring (*d* = .06) statements (see Figure 1).

[Insert Figure 1 here]

Significant main effects for group ($F_{1,16}$ = 5.16, p = .04, η_p^2 = .24) and verbal statement type ($F_{3,48}$ = 31.06, p < .001, η_p^2 = .66) were observed (Table 2). Skilled participants made significantly more verbal statements compared to their less skilled counterparts. Pairwise comparisons demonstrated that all participants made more evaluation and deep planning than monitoring and prediction statements (p < .001). There were no significant differences when comparing the number of monitoring and prediction and evaluation and deep planning statements (p >.05).

[Insert Table 2 here]

Option-generation Data

Number of options generated. ANOVA revealed a significant main effect for group ($F_{1,16}$ = 5.88, p = .03, η_p^2 = .27). Skilled participants made significantly more option statements (*M* = 5.67, *SD* = 2.03) than less skilled participants (*M* = 4.00, *SD* = 1.19). There was no main effect for condition ($F_{1,16}$ = 1.2, p = .29, η_p^2 = .07) and no Group × Context Condition interaction ($F_{1,16}$ = .3, p = .59, η_p^2 = .02).

Initial serial position of final option. A significant main effect for group was observed ($F_{1,16}$ = 4.56, p = .049, η_p^2 = .22). Skilled participants (*M* = 0.50, *SD* = 0.34) ratio score was closer to 1 indicating that they generated their final option later in the option generation process compared to less skilled participants (*M* = 0.31, *SD* = 0.37).

There were no significant context condition ($F_{1,16}$ = 3.74, p = .071, η_p^2 = .19) effect or Group × Context Condition ($F_{1,16}$ = .89, p = .36, η_p^2 = .89) interaction.

DISCUSSION

Skilled and less skilled emergency physicians managed and diagnosed simulated patient conditions. We employed a novel approach by not providing contextual information during simulated medical scenarios. We predicted that skilled participants would demonstrate superior performance and less sensitivity indicated by a less marked decrement in diagnostic accuracy when contextual information was restricted compared to less skilled participants.

Skilled participants outperformed the less skilled participants on diagnostic accuracy. All participants demonstrated lower diagnostic accuracy when context was restricted, however, skilled participants showed a lower percentage decrement. Skilled participants achieved 100% accuracy during the high- context condition. These findings extend previous work in medicine [15] by highlighting the importance of contextual information during a dynamic and evolving emergency medicine scenario.

Concurrent verbal reports were collected to examine skill-based differences in decision-making and how this may alter as a function of restricting contextual information. We predicted that skilled participants' verbal statements would be characterised by more evaluation, prediction and deep planning statements. In addition, LTWM retrieval structures remain active and allow additional contextual information to be updated when it becomes available [11]. Therefore, we expected skilled participants to make fewer evaluation, prediction and deep planning statements during the low-context scenario.

Skilled participants made more verbal statements in total compared to less skilled participants. Skilled participants made more evaluation, prediction and deep planning statements. Skilled participants were able to engage in systematic deep planning based on the evaluation of the current situation and predicted future events that could occur as the scenario progressed. Our findings are consistent with previous work on skill-based differences in distribution of verbal statement type and adds further support for LTWM theory [4,7,9,14,26]. We postulate that skilled participants' superior diagnostic accuracy is underpinned by the development of retrieval structures and associated domain-specific knowledge. Their diagnostic accuracy appeared less sensitive to reduced contextual information compared with the less skilled group, suggesting that information previously stored during earlier similar clinical encounters is able to be retrieved in the absence of current information. Previous knowledge acquired by exposure to real-world medical scenarios allow skilled participants to better access and predict future events, enabling them to plan appropriate responses in an effective and efficient manner. The less skilled participants lack of domain knowledge results in cognitive processes that primarily support monitoring and evaluation of the current situation.

Skilled participants' concurrent verbal reports did not alter across the low- and high-context conditions. However, the differences in diagnosis accuracy indicate that different cognitive processing should precede the resulting behaviour. Our current analysis does not provide information on the sequence of thoughts or how critical the verbalised cognition was to the successful outcome of patient treatment and diagnosis.

We examined skill-based differences in cognitive option generation strategies and the effect of contextual information on these processes. Within LTWM theory, retrieval structures remain actively linked to associated domain knowledge, and increase working memory capacity [9]. Therefore, skilled participants should generate more options and select the best option irrelevant of its sequence position compared to less skilled participants.

Skilled participants generated a higher number of options during the medical scenarios, selected the best option later in the sequence of options while maintaining superior diagnostic accuracy compared to less skilled participants. These data corroborate previous studies [4,7,9,27,28] reinforcing the view that skilled performance is characterised by the development of LTWM skills that allow individuals to build detailed cognitive representation knowledge of the situation.

In summary, contextual information was removed during high fidelity emergency medicine scenarios to examine its impact on diagnostic accuracy. Furthermore, skill-based differences in cognitive processes and option generation strategies that support performance on the task were examined. Skilled participants demonstrated superior diagnosis accuracy and employed a more elaborate domain-specific cognitive representation of the scenario, compared to less skilled participants. This more elaborate knowledge allowed them to generate more options and select the best quality option decision, irrelevant of where it was generated in the sequence of options. Skilled emergency physicians develop LTWM skill that allows them to use a stable cognitive processing strategy to meet the demands of the scenario irrelevant of the available contextual information. In future, researchers should develop more context-appreciative encoding categories to measure the cognitive quality of the verbal report in relation to successful performance on the task [29].

IMPLICATIONS FOR QUALITY, SAFETY AND TRAINING

Our findings have implications for the training and evaluation of decision-making skills in emergency doctors using simulation-training programmes with potential for improved patient care. First, contextual information influences diagnosis accuracy and should be carefully controlled in the scripting and stage management of scenarios, particularly those intended for ‘high stakes’ training and assessment. The performance of less skilled physicians, such as trainees or those with limited recent clinical experience, may be more sensitive to the availability of contextually relevant information. Therefore, educators and supervisors should be attuned to the learner’s level of competency, exposure and experience. Simulation scenarios are generally scalable in respect to difficulty and complexity and can be modified for educational and logistic reasons. The level of contextual information should be factored in when appraising the difficulty of scenarios, the required preparation, pre-briefing and debriefing.

Second, measurement of verbal reporting and option generation present opportunities to design education programmes that specifically target clinical reasoning and improve validity of assessment methods. Generally, immersive simulation techniques involve uninterrupted real-time cases with post hoc analysis in the form of a reflective debrief. Introducing techniques such as ‘think aloud’ verbal reporting and option generation into the scenario activity allows less skilled learners to actively reflect upon their thought processes post scenario. In conjunction with the educator, the learner can identify knowledge gaps or re-frame clinical reasoning (i.e. cognitive processing).

Third, verbal report categories presented in this study could be further developed and incorporated into educational programmes. For example, accelerated educational programmes should include practical demonstrations by experts and ‘deliberate practice’ by learners in generating diagnostic hypotheses, testing and revising hypotheses and generating alternative options. Finally, verbal reports can be used to develop cognitive processing and monitor changes in processing pre- and post an educational intervention (i.e., simulation training).

Table captions

Table 1. Verbal report statement categories and definitions

Category	**Definition**
Monitor	Statements representing information that was present or previously present in the current environment.
Evaluation	Some form of positive or negative assessment of a prior action, event or statement.
Prediction	Statements about what could, would and should occur next in the environment.
Deep Planning	Statements are about future actions and options in a future situation.

Table 2. Mean (& *SD*) number of verbal statements for group and verbal statement type

		Mean	*SD*
Group			
	Skilled	18.79	7.64
	Less skilled	15.28	6.87
Verbal Statement Type			
	Monitor	13.06	3.56
	Evaluation	22.97	8.97
	Prediction	11.64	4.44
	Deep Planning	20.47	4.75

Figure captions

Figure 1. Number of verbal statement types (with SD bars) for skilled and less-skilled participants.

REFERENCES

1 Statistics ABO. Causes of death, Australia. *Australian Bureau of Statistics*; 2008.

2 Kaufman D. Expertise in the clinical interview: a cognitive analysis. *Resident Medical Education*. 1988;27:167–72.

3 Williams AM, Ford PR, Eccles DW, et al. Perceptual-cognitive expertise in sport and its acquisition: Implications for applied cognitive psychology. *Appl. Cognit. Psychol.* 2010;25(3):432–42.

4 Ward P, Suss J, Eccles DW, et al. Skill-based differences in option generation in a complex task: a verbal protocol analysis. *Cogn Process*. 2011;12(3):289–300.

5 Cormier EM, Pickett-Hauber R, Whyte J. Cognitions and clinical performance: a comparison of high and low performing baccalaureate nursing students. *Int J Nurs Educ Scholarsh*. 2010;7(1):1-18.

6 Crognier L, Féry Y-A. Effect of tactical initiative on predicting passing shots in tennis. *Appl. Cognit. Psychol*. 2005;19(5):637–49.

7 McRobert AP, Ward P, Eccles DW, et al. The effect of manipulating context-specific information on perceptual-cognitive processes during a simulated anticipation task. *Brit J Psychol*. 2011;102(3):519–34.

8 Macmahon C, Starkes J. Contextual influences on baseball ball-strike decisions in umpires, players, and controls. *J. of Sports Sc*. 2008;26(7):751–60.

9 Ericsson K, Kintsch W. Long-term working memory. *Psychol rev*. 1995;102(2):211–44.

10 Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychol rev*. 1993;100(3):363.

11 Kintsch W. Comprehension: A Paradigm for Cognition. Cambridge, UK: Cambridge University Press; 1998.

12 Ericsson KA, Patel V, Kintsch W. How experts' adaptations to representative task demands account for the expertise effect in memory recall: comment on Vicente and Wang (1998). *Psychol rev*. 2000;107(3):578–92.

13 McRobert AP, Williams AM, Ward P, et al. Tracing the process of expertise in a simulated anticipation task. *Ergonomics*. 2009;52(4):474–83.

14 North JS, Ward P, Ericsson A, et al. Mechanisms underlying skilled anticipation and recognition in a dynamic and temporally constrained domain. *Memory*. 2011;19(2):155–68.

15 Verkoeijen PPJL, Rikers RMJP, Schmidt HG, et al. Case representation by medical experts, intermediates and novices for laboratory data presented with or

without a clinical context. *Med Educ*. 2004;38(6):617–27.

16 Paull G, Glencross D. Expert perception and decision making in baseball. *Int J Sport Psychol*. 1997;28(1):35–56.

17 Gordon JA, Tancredi DN, Binder WD, et al. Assessment of a clinical performance evaluation tool for use in a simulator-based testing environment: a pilot study. *Acad Med*. 2003;78(10):S45.

18 Issenberg BS, Mcgaghie WC, Petrusa ER, et al. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach*. 2005;27(1):10–28

19 Whyte J, Ward P, Eccles DW. The relationship between knowledge and clinical performance in novice and experienced critical care nurses. *Heart Lung*. 2009;38(6):517–25.

20 Ericsson KA, Williams AM. Capturing naturally occurring superior performance in the laboratory: translational research on expert performance. *J Exp Psychol-Appl*. 2007;13(3):115–23.

21 Ericsson KA, Kirk E. *Instructions for giving retrospective verbal reports*. Unpublished manuscript, Department of Psychology, Florida State University, Tallahassee, Florida, USA; 2001.

22 Eccles DW. Verbal reports of cognitive processes. Tenenbaum G, Eklund RC, Kamata A, editors. Measurement in Sport and Exercise Psychology. Champaign, IL: Human Kinetics; 2012. p. 103–17.

23 Whyte J IV, Cormier E, Pickett-Hauber R. Cognitions associated with nurse performance: A comparison of concurrent and retrospective verbal reports of nurse performance in a simulated task environment. *Int J Nurs Stud*. 2010;47(4):446–51.

24 Ericsson KA, Simon HA. Protocol Analysis: Verbal Reports as Data, revised edition. A Bradford Book; 1993.

25 Thomas J, Nelson J, Silverman S. Research Methods in Physical Activity. Champaign, IL: Human Kinetics; 2010.

26 Roca A, Ford PR, McRobert AP, et al. Identifying the processes underpinning anticipation and decision-making in a dynamic time-constrained task. *Cogn Process*. 2011;12(3):301–10.

27 McPherson S. The influence of player experience on problem solving during batting preparation in baseball. *J Sport Exercise Psy*. 1993;15:304–4.

28 McPherson SL, MacMahon C. How baseball players prepare to bat: Tactical knowledge as a mediator of expert performance in baseball. *J Sport Exercise Psy*. 2008;30(6):755–78.

29 Ching Yang S. Reconceptualizing think-aloud methodology: Refining the encoding and categorizing techniques via contextualized perspectives. *Comput hum*

behav. 2003;19(1):95–115.

Acknowledgements

The authors would like to thank Mrs Stephanie O'Regan, Mr Christopher Carpenter, Dr Femke Geijsel and Dr Henry Curtis for their time and knowledge that made the data collection possible.

Funding statement

The research received no funding.

Competing interests

There are no competing interests for publication.

Contribution statement

Prof. Mark Williams, Dr. John Vassiliadis and Dr. Allistair McRobert were responsible for the conception and design of the current study. Secondly, Dr. Joe Causer, Dr. John Vassiliadis, Dr. James Kwan and Dr. Allistair McRobert collected the data and engaged in many hours of subsequent analysis and interpretation of this data. Thirdly, Dr. Allistair McRobert was responsible for the initial draft of the article, revisions and approval of the final version submitted. Finally, Prof. Mark Williams, Dr. John Vassiliadis, Dr. Joe Causer and Dr. Leonie Watterson were responsible for critically revising the drafts for intellectual content and quality of the submission.