Use of head camera-cued recall and debrief to externalise expertise: a systematic review of literature from multiple fields of practice

Vivienne Isabella Blackhall,*,†,1,2 Kenneth Grant Walker,*,1,2 Iya Whiteley,‡,3 Philip Wilson§,4

ABSTRACT

Background  The study of decision making in complex naturalistic environments poses several challenges. In response to these, video-stimulated cued-recall-debrief was developed. It involves an individual wearing a head-mounted camera which records a task from their point of view. Afterwards, footage captured is reviewed along with a facilitated debrief to help externalise cognitive processes. In theory, motion, audio and visual cues generate a high level of experiential immersion which helps the expert to articulate previously hidden thoughts and actions.

Objective  To examine the current evidence for video-stimulated cued-recall-debrief as a means of explicating expert thoughts and feelings in complex tasks in a range of environments.

Study selection  MEDLINE, EMBASE, Education Resources Information Center, SPORTDiscus, PsyCINFO and Google Scholar were searched for articles containing the key terms ‘cued-recall (debrief)’, ‘decision making’, ‘skills’ and ‘video recording’. Studies were included if they examined the following outcomes: (1) feasibility, (2) extent of experiential immersion, (3) ability to generate unique insight into decision-making processes and (4) current applications. 1831 articles were identified initially, and 9 studies were included in the final review.

Findings  Video-stimulated cued-recall-debrief is associated with a high level of experiential immersion and generates between two and four times the number of recollections compared with free recall. It can be used to build models of cognitive activity and to characterise the way in which more and less skilled individuals tend to think and feel.

Conclusions  The technique could be used to explicate expertise within medicine: these insights into performance could be used as a training tool for other practitioners.

Trial registration number  CRD42017057484.

INTRODUCTION

Naturalistic decision making is that which occurs out with the laboratory setting in complex (time-critical, high-stakes, dynamic) situations. The study of such decision-making processes poses several problems for the researcher.2 While concurrent reporting can interrupt the flow of a procedure and alter the cognitive processes being studied, retrospective reporting, if not cued by real events, can be associated with bias and omission of key information. In response to these limitations, a technique known as video-stimulated cued-recall-debrief was developed by Omodei et al.2

METHODS

Data collection

Our systematic review protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO) on 6 March 2017 (trial registration number CRD42017057484).

A systematic literature search was conducted in May 2017 with the assistance of one librarian at our institution. Key terms were searched for in MEDLINE, EMBASE, Education Resources Information Center, SPORTDiscus, PsyCINFO and Google Scholar from first records to May 2017. The literature search was limited to the English
Papers were included if the participants were considered to be experts (by the authors’ own definition), undertaking any complex task (time-critical, high-stakes, dynamic conditions). No limit was placed on the type of task and could, for example, include playing a tennis match. The search included all studies describing video-stimulated cued-recall-debrief, or indeed any other similar feedback process of another name which is designed to elicit detailed recall in order to explicate the expert’s thoughts and feelings after the task. Comparators were all or none.

Studies were included if they examined one or more of the following features of the technique:

- Feasibility.
- Extent of experiential immersion.
- Ability to generate unique insight into the performer’s decision-making processes.
- Current applications.

Quality assessment

All studies had a qualitative component. To assess the quality of the papers, Critical Appraisal Skills Programme (CASP) checklists were used. There are no other published reporting guidelines which include a tool for assessing qualitative research in medical education. Manuscripts were not graded in terms of their quality, as the CASP checklist guidance notes state the checklists are designed to be used as educational tools, not scoring systems. The results were tabulated and areas of concern highlighted using a colour coding system (table 2). For mixed-methods studies, quality assessment of quantitative components was included in a separate column.

RESULTS

Nine studies were included in the synthesis, the features of which are outlined in table 3. There was significant heterogeneity in terms of the participants, which included clinicians (occupational therapists and emergency physicians) and others (orienteers, computer war games players, fire station officers and military pilots). There were major differences in how an ‘expert’ was defined. One study stipulated satisfaction of a 6-point checklist in order to qualify, which included having been nominated by a hierarchical superior and holding the highest clinical hospital grade. The checklist did not include years of experience, which is known to be a poor surrogate marker for expert performance. Despite this, many papers used this to define their expert population.
Table 2: Quality analysis

<table>
<thead>
<tr>
<th></th>
<th>Unsworth²⁷</th>
<th>Unsworth⁶</th>
<th>Unsworth⁹</th>
<th>Pelaccia et⁷</th>
<th>Omodei¹⁹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear statement of aims</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Qualitative methodology</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Research design appropriate?</td>
<td>Yes.</td>
<td>No.</td>
<td>No.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Recruitment strategy appropriate?</td>
<td></td>
<td>Unclear.</td>
<td>Unclear.</td>
<td></td>
<td>Unclear</td>
</tr>
<tr>
<td>Ethical issues considered?</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Rigorous data analysis?</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Quantitative methodology</td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
</tr>
</tbody>
</table>

**Table 2 Continued**

<table>
<thead>
<tr>
<th></th>
<th>Omodei¹¹</th>
<th>Omodei and McLennan³⁸</th>
<th>Elliot et⁸</th>
<th>Solodilova-Whiteley and Johnson¹²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear statement of aims</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Qualitative methodology</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Research design appropriate?</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Recruitment strategy appropriate?</td>
<td></td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Ethical issues considered?</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
<tr>
<td>Relationship between researcher and participants considered?</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
<td>Yes.</td>
</tr>
</tbody>
</table>
Systematic review

There were small variations in the way in which the intervention was described. True to the original description, all participants wore a head-mounted camera while undertaking a complex task and subsequently undertook a facilitated video-stimulated ‘cued-recall-debrief’ process to generate a commentary of the episode. The majority of studies used the authentic ‘non-directive’ approach to the debrief; however, one used a novel, semi-structured format. In five of the nine studies, the individual facilitating the debrief was poorly defined, referred to as ‘an experimenter or member of the team’, and there was no description in any of the studies relating to their level of experience or training.

All studies were observational and had a qualitative component. Four were purely qualitative, and the remainder used a mixed-methods approach. In the majority of the studies, the qualitative component involved transcription and coding of the cued-recall-debrief commentary. Two coding approaches were described: a priori or emergent. Generally, where a relevant cognitive model existed in the literature, coding proceeded in an a priori manner. Sometimes, this generated new emergent codes which did not fit with the existing framework. Other studies adopted a wholly emergent code approach to coding. Irrespective of how they were generated, emergent codes were used to perform a thematic (or similar) analysis and gain new understanding of cognitive processes or generate novel theoretical models.

Three methods of quantitative analysis were described in the literature:

- Counting the frequency of codes encountered.
- Using a computer program to generate quantity of codes as a function of time.
- A questionnaire measuring participant perceptions of the utility of the intervention in terms of immersion and ability to trigger recollections and insights.

Quantitative methods were employed when some form of comparison was taking place:

- Comparing cognitive activity of different participant groups (eg, novice and expert).
- Comparing level of immersion and recollections generated from cued-recall-debrief with other techniques (eg, free recall).
- Comparing the content of recollections from cued-recall-debrief with other established cognitive models.
- Stratifying the level of performance through the use of rating scales.

The studies could be categorised into three major groups depending on their aims:

- Basic description of the technique with outcome data.
- Comparing the effectiveness of cued-recall-debrief with other techniques in studying naturalistic decision making.
- Applications of the information generated from cued-recall-debrief:
  - Generation of novel insight or theory into cognitive processes.
  - Comparing the cognitive processes of novices and experts or good and poor performers.

Each of these categories will be considered in turn.

Basic descriptive study

A case report by Unsworth demonstrated that occupational therapists and their clients felt that the use of head-mounted camera during their interaction was acceptable. The equipment was portable, did not affect therapist mobility nor interfere with...
<table>
<thead>
<tr>
<th>Paper</th>
<th>Omodei and McLennan 8</th>
<th>Elliot 9</th>
<th>Unsworth 4</th>
<th>Unsworth 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Publication year</strong></td>
<td>1994</td>
<td>2000</td>
<td>2001</td>
<td>2001</td>
</tr>
<tr>
<td><strong>Participants (n)</strong></td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>3 (+2 novice)</td>
</tr>
<tr>
<td><strong>Participant type</strong></td>
<td>Orienteers</td>
<td>Computer war games players</td>
<td>Occupational therapists</td>
<td>Occupational therapists</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>3 male:3 female</td>
<td>All male</td>
<td>All female</td>
<td>All female</td>
</tr>
<tr>
<td><strong>Mean age</strong></td>
<td>16</td>
<td>Not stated.</td>
<td>Not stated.</td>
<td>Not stated.</td>
</tr>
<tr>
<td><strong>Aim</strong></td>
<td>Compare two methods for studying naturalistic decision making (NDM).</td>
<td>Compare two methods of studying NDM.</td>
<td>Basic description of technique</td>
<td>Compare cognitions of experts and novices.</td>
</tr>
<tr>
<td><strong>Context</strong></td>
<td>Competitive orienteering circuits.</td>
<td>Mission commander in simulated air defence tasks.</td>
<td>An encounter between occupational therapist and client during a physical rehabilitation episode.</td>
<td>An encounter between occupational therapist and client during a physical rehabilitation episode.</td>
</tr>
<tr>
<td><strong>Comparator</strong></td>
<td>Free recall.</td>
<td>Adversarial crew cognitive walkthrough.</td>
<td>Concurrent reporting—free recall. Retrospective reporting with assisted recall (audio or external video).</td>
<td>No.</td>
</tr>
<tr>
<td><strong>Primary outcome</strong></td>
<td>Level of experiential immersion. Extent of insight into decision-making processes.</td>
<td>Extent of insight into decision-making processes.</td>
<td>Feasibility. Level experiential immersion (surrogate markers of Point of view camera and visibility of cues).</td>
<td>NA.</td>
</tr>
<tr>
<td><strong>Methods of data collection and/or analysis</strong></td>
<td>1. Orienteering review analysis checklist (quantitative). 2. Structured interview protocol comparing each technique (qualitative).</td>
<td>1. MacShapa program: quantity of information taken as a function of time (quantitative). 2. Transcription and coding of commentaries with quantitative count of recollections. 3. Informal researcher observations (qualitative).</td>
<td>Author’s own experiences.</td>
<td>NA.</td>
</tr>
<tr>
<td><strong>Secondary outcome</strong></td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
<td>Compare the cognitions of experts vs novices.</td>
</tr>
<tr>
<td><strong>Method of data collection and/or analysis</strong></td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
<td>Transcription and coding of commentaries. Quantitative count of recollections and qualitative thematic analysis.</td>
</tr>
<tr>
<td><strong>Coding approach</strong></td>
<td>NA.</td>
<td>A priori—framework not stated.</td>
<td>NA.</td>
<td>A priori (Mattingly and Fleming 14, Schell and Cervero 15 clinical reasoning frameworks).</td>
</tr>
<tr>
<td><strong>Statistical method (quantitative only)</strong></td>
<td>Wilcoxon signed-rank test comparing differences between techniques.</td>
<td>Basic descriptive, no formal statistical analysis applied.</td>
<td>NA.</td>
<td>( \chi^2 ) analysis for differences between groups. Simple descriptive statistics.</td>
</tr>
<tr>
<td><strong>Potential utility</strong></td>
<td>Training students/beginners to think like experts.</td>
<td>Training students/beginners to think like experts.</td>
<td>Training.</td>
<td>Training. Trainers can target the cognitive gap that separates novices and experts.</td>
</tr>
</tbody>
</table>

Continued
### Table 3: Continued

<table>
<thead>
<tr>
<th>Paper</th>
<th>Unsworth⁵</th>
<th>Omodei¹⁰</th>
<th>Omodei¹¹</th>
<th>Solodilova-Whiteley and Johnson¹²</th>
<th>Pelaccia et al¹³</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Publication year</strong></td>
<td>2005</td>
<td>2005</td>
<td>2006</td>
<td>2014</td>
<td></td>
</tr>
<tr>
<td><strong>Participants (n)</strong></td>
<td>13</td>
<td>9</td>
<td>12</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td><strong>Participant type</strong></td>
<td>Occupational therapists</td>
<td>Fire station officers</td>
<td>Fire station officers</td>
<td>Military pilots</td>
<td>Emergency physicians</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>All female</td>
<td>All male</td>
<td>All male</td>
<td>All male</td>
<td>11 male, 4 female</td>
</tr>
<tr>
<td><strong>Mean age</strong></td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>42</td>
</tr>
<tr>
<td><strong>Design</strong></td>
<td>Qualitative</td>
<td>Mixed methods</td>
<td>Mixed methods</td>
<td>Qualitative</td>
<td>Qualitative</td>
</tr>
<tr>
<td><strong>Aim</strong></td>
<td>Conceptualise the clinical reasoning of occupational therapists.</td>
<td>Conceptualise the decision making of firefighters.</td>
<td>Compare the cognitions of good vs poor decision makers.</td>
<td>Conceptualise information use by pilots.</td>
<td>Conceptualise the clinical reasoning of emergency physicians.</td>
</tr>
<tr>
<td><strong>Context</strong></td>
<td>Client encounter during physical rehabilitation.</td>
<td>Simulated emergency fire incident.</td>
<td>Simulated emergency fire incident.</td>
<td>Simulated flight.</td>
<td>Encounter with patient with life-threatening condition.</td>
</tr>
<tr>
<td><strong>Comparator</strong></td>
<td>No.</td>
<td>No.</td>
<td>No.</td>
<td>No.</td>
<td>No.</td>
</tr>
<tr>
<td><strong>Primary outcome</strong></td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
</tr>
<tr>
<td><strong>Methods data collection/analysis</strong></td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
<td>NA.</td>
</tr>
<tr>
<td><strong>Secondary outcome—application</strong></td>
<td>Generate additional insight into clinical reasoning of occupational therapists.</td>
<td>Generate additional insight into how cognitive activity is used in firefighters.</td>
<td>Evaluate differences in cognitive content between poor and good decision makers.</td>
<td>Generate a model demonstrating information gathering and assimilation in a cockpit.</td>
<td>Define the cognitive strategies of emergency physicians in making a clinical diagnosis.</td>
</tr>
<tr>
<td><strong>Method of data collection and/or analysis</strong></td>
<td>Transcription and coding of CRD commentary.</td>
<td>Transcription and coding of CRD commentary with quantitative count of recollection types. Pattern compared with Klein’s recognition-primed decision model.</td>
<td>Transcription and coding CRD commentary. Quantitative count of type of recollections and qualitative thematic analysis. 10-point rating scale to rate and stratify performance.</td>
<td>Transcription and coding of CRD commentary in order to generate a model of information flow.</td>
<td>Transcription and coding of CRD commentary with thematic analysis.</td>
</tr>
<tr>
<td><strong>Coding approach (framework used)</strong></td>
<td>A priori (Mattingly and Fleming, Schell and Cervero clinical reasoning frameworks).</td>
<td>A priori (Cognitive Process Tracing Categorisation Scheme (CPTCS)).</td>
<td>A priori (CPTCS).</td>
<td>Emergent.</td>
<td>Emergent.</td>
</tr>
<tr>
<td><strong>Statistical method (quantitative only)</strong></td>
<td>NA.</td>
<td>NA.</td>
<td>χ² analysis comparing differences between groups.</td>
<td>NA.</td>
<td>NA.</td>
</tr>
<tr>
<td><strong>Potential utility</strong></td>
<td>Training students/beginners to think like experts.</td>
<td>Training.</td>
<td>Training. Trainers can target the cognitive gap that separates novices and experts.</td>
<td>Implications for human/systems interface design.</td>
<td>Training: Trainers can target the cognitive gap that separates novices and experts.</td>
</tr>
</tbody>
</table>

CRD, cued-recall-debrief; NA, not applicable.
the therapist–client interaction: participants ‘forgot’ about the camera’s presence and were able to ‘continue as usual’.

Unsworth was able to generate accurate ‘own point-of-view’ footage, which displayed the rich visual cues of the client’s facial expressions (known to be important in generating immersion). No comment was made as to the extent of experiential immersion experienced by the therapist nor the extent of the recollections generated.4

The main criticism of Unsworth’s paper is that the outcomes were poorly described and interlaced with anecdotal findings from other papers, making identification of novel data challenging.

**Comparison with other techniques**

Two studies compared the effectiveness of cued-recall-debrief with other methods of studying naturalistic decision making (free recall and a ‘walkthrough’ technique).8,9 Effectiveness was measured in terms of the extent of experiential immersion and verbalised recollections.

Omodei and McLennan compared cued-recall-debrief with free recall, in a group of orienteers.8

Cued-recall-debrief was associated with increased experiential immersion (4.4 vs 6.33, p<0.01) and allowed the individual to recount between two and four times the amount of detail compared with free recall. Cued-recall-debrief generated more recollections related to thoughts (5.3 vs 6.87, p<0.01) and feelings (4.42 vs 6.68, p<0.01). All participants stated that cued-recall-debrief was the superior method in terms of recalling mental processes.

Elliott et al compared the effectiveness of cued-recall-debrief with a cognitive walkthrough model in participants undertaking a simulated air defence task.9 The walkthrough model (adversarial crew cognitive walkthrough or ACCW) involved pausing game play when a significant episode occurred in order for researchers to pose questions about the expert’s thoughts and feelings.

The paper concluded that ACCW was more time-efficient. ACCW took approximately half the time of cued-recall-debrief in generating the same quantity of recollections (1 hour 42 min vs 3 hours 28 min).

Compared with ACCW, cued-recall-debrief generated between two and three times the number of recollections related to the expert’s own thoughts. The techniques were comparable in their ability to generate recollections relating to their adversary. These results were demonstrated graphically with no specific numerical values for each type of recollection. Furthermore, no formal statistical analysis was undertaken to determine whether the difference between the groups was statistically significant. Experts stated that both techniques were useful and provided rich insights into their cognitive processes. Interestingly, the authors concluded that ACCW was the more effective technique, presumably because it was more time-efficient.9

**Generation of new theory**

Unsworth used two established clinical reasoning models14,15 as a priori frameworks in order to map the clinical reasoning of occupational therapists.5 She was not only able to demonstrate the most commonly used types of reasoning (procedural, interactive and conditional) and the way in which they were used (in rapid succession and simultaneously), but she also identified an emergent subcategory of reasoning (generalisation reasoning).

Solodilova-Whiteley and Johnson uncovered the cognitive processes of military pilots.12 Through an inductive, evolutionary data analysis technique, they generated an eight-stage model of how pilots acquire and use information which could be used to help inform the design of more efficient cockpit interfaces.

The authors found that pilots used previous experience to generate a model outlining the expected course of the flight which was articulated in a preflight brief. Key stages in the flight were bookmarked with expected values, which were compared with actual values. Pilots gathered information from many sources, at two-minute intervals, and values were always interpreted relative to one another. Information gathering was supported by common information structures such as checklists or standard operating procedures. On completion of information gathering, pilots anticipated potential problems. A process of problem solving and information organisation followed before pilots reached a ‘point of clarity’ and executed their decision. This entire process was converted to knowledge and experience, banked for future flights.

Solodilova-Whiteley and Johnson’s paper gives us some general insights into expert decision-making processes. First, experts rely on previous experience to pattern-match situations. Second, they demonstrate good situational awareness: systematically gathering and assimilating data, while anticipating future events. Lastly, they use strategies to reduce the burden on their working memory. Doing things in the same way every time, for example, using checklists, helps to reduce the burden on working memory and ensures data gathering is systematic.12

**Generation of additional insight**

Omodei et al studied the decision-making processes of firefighters using cued-recall-debrief.7 The information gathered by her team was compared with an established decision-making model (recognition-primed decision making).16 Some similarities were demonstrated between the data gathered by cued-recall-debrief and the recognition-primed decision model (decision making is primarily intuitive); however, the former revealed novel information: 20% of cognitive activity was devoted to affect and motivation. This led the researchers to postulate that individuals who are more effective in terms of self-monitoring and emotional regulation may be more skilled performers. This hypothesis was confirmed in their follow-on paper,10 which will be considered later.

Pelaccia et al studied emergency department physicians during a patient encounter.6 The following insights into the decision-making processes of participants were revealed:

- All generated at least four hypotheses.
- Most hypotheses were generated fast and without conscious effort.
- Physicians remained vigilant to the emergence of new hypotheses (ie, avoided premature closure).
- Hypotheses were rank-ordered and the order was not consistent with standard probability theory.

**Applications of cued-recall-debrief**

These studies did not evaluate the feasibility or effectiveness of cued-recall-debrief as a means of studying naturalistic decision making; these qualities were assumed. These studies had three distinct aims:

- Generation of additional insight or novel theory into the cognitive processes of the expert.5,7,10,12
- Comparing the cognitive processes of novices and experts.6
- Comparing cognitive processes in good versus poor performance.11
Comparison of novice and expert performance

Unsworth demonstrated important differences in the clinical reasoning of expert and novice occupational therapists. While experts demonstrated a greater capacity to reason interactively and conditionally with their clients, novices tended to demonstrate a more procedural, textbook approach. In addition, the study showed that experts were more likely to:

- Use pattern recognition techniques.
- ‘Think on their feet’ (improvisation).
- Use more than one type of reasoning at once (multitasking).

Comparing good versus poor performance

Omodei and McLennan’s follow-on study involving firefighters investigated the differences in cognitive activity associated with ‘good’ versus ‘less effective’ decision making. All participants were experts and their performance was stratified into poor or good by several fire officers using a 10-point scale. The performance standard was peer-referenced: the six best scores were deemed to represent good performance and the six lowest to equate with poor performance.

There was a significant difference in the way in which cognitive activity was used. Poorer performers tended to report a greater frequency of thoughts related to self-monitoring and emotional regulation (X²(5; n = 485) = 27.4, p < 0.001), the content of which was uniformly negative. Good performers thought less frequently about self-monitoring and regulation, but when they did they exhibited no such negative evaluations. The authors surmised that effective decision making is associated with fewer negative self-evaluations, indicating more effective emotional regulation.

The qualitative analysis revealed that the poorer decision makers had a tendency to become overwhelmed by the scenario (cognitive overload). Consequently, the capacity of their working memory was reduced. They demonstrated poorer situational awareness and a tendency to commit decision-making pitfalls (eg, anchoring). They tended to focus on only one salient feature of the situation while ignoring emerging threats.

More skilled participants demonstrated better situation awareness as they:

- Kept track of events.
- Anticipated developments.

Good performers also reacted to new threats without undue irritation or concern. Interestingly, there was no correlation between level of performance and number of years of experience.

DISCUSSION

In summary, the literature suggests that video-stimulated cued-recall-debrief is a useful and feasible means of gaining insight into expert performance. However, these results should be interpreted with caution due to the small body of literature describing the intervention and the methodological flaws identified in many of the studies. The head-mounted camera is portable and does not appear to interfere with the episode being studied (although this latter observation is purely anecdotal). Cued-recall-debrief is associated with a high level of experiential immersion and allows an individual to recount between two and four times the number of recollections when compared with free recall. The method is particularly useful in terms of gaining insight into one’s own actions and feelings. It is however more time-intensive in terms of time than other information elicitation techniques such as cognitive walkthrough. A cognitive walkthrough is not always possible in naturalistic decision-making settings. For example, it would be dangerous to ask a surgeon to pause and discuss their thoughts and feelings when performing an emergency tracheostomy.

The majority of the papers included in this review considered the technique’s applications:

- Generating additional insight or novel theory relating to expert cognition.
- Comparing the cognitive processes of two groups of individuals undertaking the same task.

These studies demonstrate commonalities in the cognitive and affective processes of experts across several disciplines. Generally, experts tended to use intuitive, as opposed to analytical, reasoning. These types of reasoning can be conceptualised using the dual process theory, which posits that two systems of decision making exist. System 1 is intuitive, automatic and fast and largely based on pattern matching. It relies on having built up a bank of previous experience. System 2 is slower and analytical and involves a degree of problem solving.

Other features demonstrated by experts were as follows:

- An ability to multitask.
- An ability to improvise.
- Awareness of and ability to use strategies to avoid common pitfalls: checklists, cognitive strategies (avoiding premature closure).
- Good situational awareness.
- Effective self-monitoring and emotional regulation.

The final point relating to emotional regulation is in keeping with what we know from some of the existing decision-making literature. Too much concern about how well one is doing in a task sometimes disrupts performance by loading short-term memory with pointless thoughts. As controlling thoughts and behaviours is one of the tasks performed by system 2, perhaps experts have a more efficient or better developed sense of analytical reasoning.

Understanding the way in which expert medical practitioners think may help to support other medical practitioners model their cognitive processes in order to develop their own expertise.

A common limitation of the studies was a lack of transparency as to the relationship of the lead researcher to the activity being studied. As a researcher, being both an ‘insider’ and ‘outsider’ can have an impact on the way in which data are collected and analysed, and therefore it is important to reflect and comment on this influence.

One paper applied no formal statistical analysis when comparing numerical values between two groups. This makes it difficult for us to surmise whether the differences between the groups were statistically significant.

For the majority of the studies, the debrief facilitator and their level of experience was poorly defined. Facilitating a debrief is a difficult skill that requires practice. As a beginner, there is a natural tendency to ask ‘why’, leading the expert to evaluate their performance, rather than to facilitate the recall of their cognitive and affective processes. To maximise the integrity and reproducibility of the technique, we should ensure that reporting is explicit in terms of the debriefer’s training experience and introduce a recommendation for core debriefing competencies to be met ahead of engaging with the technique.

A limitation of this review is the small number of studies which were included, a reflection of the small body of literature which exists in this domain. Many papers were written by the same authors, and sometimes within the same year. This may introduce an element of confirmation bias whereby researchers unintentionally focus on data which support assumptions about the technique, rather than challenging their theoretical stance. The typically small sample sizes encountered do not necessarily...
raise concerns, qualitative focuses on indepth and purposeful sampling, even single cases.20

CONCLUSIONS
To the authors’ knowledge, this is the first systematic review considering the current evidence base for video-stimulated cued-recall-debrief. The authors tentatively conclude that the technique appears to be a feasible and valid method of gaining insight into an expert’s thoughts and feelings. Additionally, it can be a useful tool in building models and patterns of cognitive activity and characterising the way in which more and less skilled individuals tend to think and feel.

Future research might include expanding the use of the technique in medicine, with attention to technical and non-technical skills. Specifically, this could examine whether there are particular environments (eg, acute medical care environments such as emergency departments, or elective treatment areas), procedures (eg, coronary artery angiography, epidural insertion or colonoscopy) or individuals for which this technique works best. The technique could also be used to determine whether the insights gained into the cognitive processes of experts are transferable to a group of learners and helps to support the development of their own expertise. Indeed, shortening the expert learning curve for medical procedures may help to improve patient safety. Lastly, the impact of the intervention on the activity being studied could be evaluated. Although inanimate technologies and materials tend to be accepted as part of the backdrop in medicine, a head-mounted camera may influence what people think, say and do.

Acknowledgements Professor Jennifer Cleland (Centre for Health Education Research and Innovation, College of Life Sciences and Medicine, University of Aberdeen) advised on the methodology. Database expertise was provided by Rob Polson, senior librarian at the Centre for Health Science, Inverness.

Collaborators Professor Jennifer Cleland (Centre for Health Education Research and Innovation, College of Life Sciences and Medicine, University of Aberdeen).

Contributors PW is the guarantor. VIB drafted the manuscript. All authors contributed to the genesis of the selection criteria. VIB and PW developed the search strategy. VIB, PW and KGW screened the papers for inclusion. All authors read, gave feedback and approved the final version of the manuscript.

Funding No funding was received for this work.

Competing interests IW authored two of the papers under review, under the names I Solodilova and I Solodilova-Whiteley. The remaining authors have no competing interests.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES