Does warm-up using mental practice improve crisis resource management performance? A simulation study†

M. A. Hayter1*, M. D. Bould2,3, M. Afsari1, N. Riem4, M. Chiu5 and S. Boet3,5

1 Department of Anesthesiology and the Li Ka Shing Knowledge Institute, Allan Waters Family Simulation Center, St. Michael’s Hospital, University of Toronto, Toronto, ON, Canada
2 Department of Anesthesiology, The University of Ottawa Skills and Simulation Centre, The Children's Hospital of Eastern Ontario, Ottawa, ON, Canada
3 The Academy for Innovation in Medical Education, University of Ottawa, Ottawa, ON, Canada
4 Department of Anaesthesia and Intensive Care Medicine, Kantonsspital Hospital, Liestal, Switzerland
5 Department of Anesthesiology, The University of Ottawa Skills and Simulation Centre, The Ottawa Hospital, Ottawa, ON, Canada

* Corresponding author. E-mail: megan.hayter@utoronto.ca

Editor’s key points
- A range of educational approaches are needed to develop skills to manage crises
- Imagining dealing with a crisis [mental practice (MP)] has been shown to improve technical skills
- This study evaluates the effect of MP on non-technical skills.
- No benefit was found using MP before a simulated clinical crisis
- Further research in this area is warranted.

Background. Mental practice (MP) is defined as the ‘symbolic rehearsal of a physical activity in the absence of any gross-muscular movements’ and has been used in sport and music to enhance performance. In healthcare, MP has been demonstrated to improve technical skill performance of surgical residents. However, its effect on crisis resource management (CRM) skills has yet to be determined. We aimed to investigate the effect of warm-up with MP on CRM skill performance during a simulated crisis scenario.

Methods. Following ethics board approval, 40 anaesthesia residents were randomized. The intervention group performed 20 min of MP of a script based on CRM principles. The control group received a 20 min didactic teaching session on an unrelated topic. Each subject then managed a simulated cardiac arrest. Two CRM experts rated the video recordings of each performance using the previously validated Ottawa GRS. The time to start chest compressions, administer epinephrine, and give blood was recorded.

Results. There was no significant difference between the intervention and control groups: total Ottawa GRS score was 24.50 (18.63–28.88 [6.50–34.50]) (median (inter-quartile range [range])) vs 20.50 (13.00–29.13 [6.50–34.50]) (P=0.53); the time to start chest compressions 146.0 s (138.0–231.0 [115.0–323.0]) vs 162.5 s (138.0–231.0 [100.0–460.0]) (P=0.27), the time to epinephrine administration 163.0 s (151.0–187.0 [111.0–337.0]) vs 187.0 s (164.0–244.0 [115.0–310.0]) (P=0.09), and the time to blood administration 220.5 s (130.8–309.0 [92.0–485.0]) vs 252.5 (174.5–398.8 [65.0–527.0]) (P=0.48).

Conclusion. Unlike technical skills, warm-up with MP does not seem to improve CRM skills in simulated crisis scenarios.

Keywords: patient simulation; behavioural research; educational models; communication

Accepted for publication: 18 July 2012

Mental practice (MP), defined as the ‘cognitive rehearsal of a skill in the absence of an overt physical movement’, is used in various high stakes fields, such as sport and music, to warm-up before the actual performance.1 2 The effectiveness of MP has been attributed to the fact that mentally simulating an action activates the same neural representation as the actual action.3 In healthcare, warm-up with MP has been demonstrated to improve technical skills while performing various tasks.3–6 Compared with traditional book learning, MP improved novice performance of cystoscopy skills6 and had an additive effect on performance compared with physical practice alone when performing basic surgical skills and simulated cricothyrotomies.5 7 In surgical trainees who had already performed an average of 11 actual laparoscopic cholecystectomies, MP resulted in greater performance improvements than training on a part task simulator when later performing a laparoscopic cholecystectomy on the same part task trainer.8 Warm-up using a part-task trainer has been found to improve surgical techniques in the operating theatre.9 Although warm-up with MP has demonstrated to be effective in improving technical skill performance.
its effect on non-technical skill performance has yet to be investigated.

Crisis resource management (CRM) comprises a set of non-technical skills that are required to manage crises effectively such as situation awareness, decision-making, task management, and team working. Although CRM skills have widespread applicability for patient safety, learners may have limited exposure to crises during their training due to their infrequent occurrence and the propensity for consultant physicians to take over the leadership role during crises. When appropriately used, simulation provides a safe learning environment for learners to practice CRM skills, which have been shown to transfer to the clinical arena. For example, one simulated session on weaning from cardiopulmonary bypass (CPB) resulted in improved technical and non-technical skills in weaning real patients from CPB. Moreover, a single simulation session resulted in improved CRM skill performance that was retained over a 3-month period.

Although increased simulation-based practice of CRM skills may be beneficial, it is logistically challenging for all physicians to attend simulation sessions to maintain their CRM skills. MP thus may have a role in the maintenance of these skills. Currently, the effect of warm-up using MP on CRM skill performance is unknown. We hypothesized that warm-up using MP would improve the CRM performance during a simulated cardiac arrest in anaesthesia trainees.

Methods
Participation
After institutional review board approval (St Michael’s Hospital, Toronto) and informed consent, 40 postgraduate anaesthesia trainees years 1 to 5 (University of Toronto, Canada) were randomized to one of the two groups.

Study design and intervention
Both groups received a standardized orientation to the simulator. Participants in the intervention group (MP) participated in 20 min quiet mental rehearsal of a script based on key CRM principles (Supplementary Appendix 1). Participants in the MP group were instructed to imagine a clinical crisis while reviewing the script. The time commitment of 20 min was based on a meta-analysis that determined that 20 min was the optimal duration for MP for both cognitive and physical tasks. Participants in the control group (no MP) received 20 min of a didactic teaching session on an unrelated subject (technical aspects of the management of intra-operative high airway pressure), excluding any aspect of CRM. Participants then completed the Mental Imagery Questionnaire (MIQ) after completing either 20 min of MP or 20 min of didactic teaching. All participants then had to manage a standardized crisis simulated scenario (Fig. 1).

MP script development (Supplementary Appendix 1)
The MP script is a text document that guides subjects through a process of anaesthesia CRM in preparation for managing a crisis. This script was developed using a qualitative ‘cognitive walk through’ technique. Two staff anaesthetists with experience in CRM were interviewed after a pilot simulation scenario and these interviews were analysed to determine key points for the CRM MP script. The two anaesthesiologists performed the same simulated resuscitation as the study scenario. Both performances were video recorded. After the scenario, each anaesthesiologist was interviewed using a ‘cognitive walkthrough’ for the procedure. This approach is similar to the verbal protocol analysis used extensively in research on expertise, except that in our case, the anaesthesiologist did not actually perform the task while being interviewed. To prevent interference with the time critical processes of CRM, the experts were asked to describe the steps of the CRM performance as they had experienced them while watching the video of their performance. While watching the video, they were asked to report the visual, cognitive, and kinaesthetic haptic cues they had experienced throughout the scenario. A researcher guided the participants through the mental imagery process, prompting for cues where necessary to ensure detail. Each interview was audiotaped. Finally, the cues for CRM from both interviews were merged to create a single MP script. The cues were colour coded based on the four categories of the Anaesthesia Non-Technical Skill...
A graphical display of the patient's vital signs through invasive blood pressure cuff, ECG, pulse oximetry, end-tidal (CO₂). If asked, participants were told that the platelet count was 66 000 litre⁻¹. All scenarios were video recorded.

Ninety seconds after the consultant anaesthetist left, the surgeon began suctioning blood from the simulated surgical site, resulting in a haemodynamically unstable patient. When the surgeon finally had control of the bleeding 2 min later, 2000 ml of simulated blood was in the suction canister by which time the patient was in bradycardic PEA. Following at least two cycles of cardiopulmonary resuscitation and the administration of one unit of blood, the mannequin returned to sinus rhythm and the scenario was terminated. Following four cycles of cardio-pulmonary resuscitation, if the participant did not request blood, the circulating nurse (a confederate) directed the participant’s attention to the 2000 ml of blood in the suction canister. The scenario was terminated with the administration of one unit of packed red blood cells. None of the participants failed to administer blood. The scenario duration was an average of 10 min.

A defibrillator, or blood products, or both were brought to the room upon the participant’s request. Confederate team members followed the instructions and orders of the participant but did not suggest any actions. All simulated scenarios were completed in a simulated theatre environment using a high-fidelity mannequin (SimMan®; Laerdal, Toronto, ON, Canada) with standard non-invasive monitors applied (non-invasive blood pressure cuff, ECG, pulse oximetry, end-tidal CO₂). A graphical display of the patient’s vital signs throughout the session was recorded and overlaid on the video footage.

Measurement instruments and outcomes

The primary endpoint of the study was the effect of warm-up using MP on the CRM performance. The performance was assessed using video recordings and the Ottawa Global Rating Scale (GRS) by two independent raters blinded to group allocation (Supplementary Appendix 2) after a rater training session. The GRS is a validated assessment tool for the measurement of CRM. It is composed of five CRM skill categories that were derived from the CRM literature—communication, leadership, resource utilization, problem solving, and situation awareness. Each category is composed of a seven-point Likert scale with alternating anchored descriptors. In addition to the five skill categories, there is an overall CRM performance assessment score. Because the MP script was developed using the previously validated ANTS scale, we opted to evaluate the participants using the validated Ottawa GRS in order to avoid the chance of simply testing memorization of the script and to attempt to measure more general principles of CRM. A Mann–Whitney test was used to compare the performance of both groups.

Secondary endpoints include:

(i) The effect of warm-up using MP on time to perform resuscitation tasks: time to starting chest compressions, time to epinephrine administration, and time to blood administration as measured by blinded video review.

(ii) The effect of warm-up using MP on the perception of readiness for CRM performance as assessed by a modified Mental Imagery Questionnaire (mMIQ, Supplementary Appendix 3 adapted from previously validated MIQ). The MIQ is a tool used to measure the richness of the individuals’ imagery process before participating in an activity. The MIQ is a seven-point Likert scale composed of eight questions and was completed by each participant just before their simulation scenario. A Spearman rank correlation coefficient was used to determine the relationship between the quality of the imagery and the performance.

Statistical analysis

The sample size was calculated a priori. The sample size was calculated to demonstrate a large effect of MP. In the field of psychology and education, a Cohen’s d effect size >0.80 is considered large and acceptable for an educational intervention. Therefore, considering an effect size of 0.8 and a two-tailed α=0.05 and a power of 0.8, we calculated that a total sample size of 34 subjects was required.

In order to adjust the sample size calculations for non-parametric analysis, we performed a Pitman asymptotic relative efficiency (ARE) to estimate the sample size, which compares the efficiency of two competing test statistics (i.e. t-test and Mann–Whitney U) and represents the ratio of sample sizes needed to achieve equal power. We assumed a priori that the underlying distribution of the Ottawa GRS scores would be normal, therefore resulting in an ARE of 0.955 and would require a sample size of 36 (34/0.955=35.6). Allowing for technical failures, we invited 40 anaesthesia trainees to participate in the study. However, a posteriori histogram of our data revealed that the distribution was not normal. As the ARE of the Mann–Whitney U is never <0.864, a more conservative estimate of required sample size would have been 40 residents (34/0.864=39.3). Although this was an oversight on our part, we did recruit 40 residents in case of technical problems with the simulations, and in fact there were no technical problems and we had useable data for all 40 residents. Therefore, our sample size was appropriate.

Subject characteristic data were analysed using the Mann–Whitney test for ordinal data and Fisher’s exact test for nominal data. The inter-rater reliability was assessed...
using the intra-class correlation coefficient for the total Ottawa GRS score between the two independent raters. A two-tailed \( P \) value of <0.05 was considered significant for all analyses. Statistical analysis was performed using SPSS 16.0 (SPSS Inc., Chicago, IL, USA).

**Results**

Baseline subject characteristics are shown in Table 1. A total of 40 anaesthesia residents were recruited. There was no significant difference between the groups with respect to postgraduate year and the number of previous simulator sessions. The inter-rater reliability of the global rating scale (GRS) was excellent with an average measures intra-class correlation coefficient of 0.88.

**Impact of MP on performance**

The median (inter-quartile range [range]) for the intervention group Ottawa GRS score 24.5 (18.6–28.9 [6.5–34.5]) was not significantly different from the control group, 20.5 (13.0–29.1 [6.5–34.5]) \( (P=0.53) \) (Fig. 2). There was no significant difference between the control 39.0 (36.0–43.0 [24.0–47.0]) and intervention group 37.6 (32.0–43.0 [28.0–45.0]) \( (P=0.53) \) on mental imagery questionnaire score (Fig. 3). There was also no significant difference in time to CPR between the intervention 146.0 s (138.0–231.0 [115.0–323.0]) and control groups 162.5 s (138.0–231.0 [100.0–460.0]) \( (P=0.27) \), time to epinephrine administration between intervention 163.0 s (151.0–187.0 [111.0–337.0]) and control groups 187.0 s (164.0–244.0 [115.0–310.0]) \( (P=0.09) \), and time to blood administration between intervention 220.5 s (130.8–309.0 [92.0–485.0]) and control groups 252.5 s (174.5–398.8 [65.0–527.0]) \( (P=0.48) \) (Fig. 4).

**Relationship between MIQ and performance**

There was no significant correlation between overall GRS performance and scores on the mMIQ as measured by Spearman’s rho \( (P=0.166) \).

**Discussion**

Contrary to our hypothesis, this study found that 20 min of MP does not significantly improve the CRM performance of anaesthesia residents’ performance of a simulated cardiac arrest. Our findings contrast with the existing literature from the surgical and sports psychology fields.8,9 The majority of published research within the fields of sport psychology and surgery has focused on the use of MP for technical tasks such as a golf swing, scoring a goal, or suturing.21

The use of MP in medical education has its roots in dual coding theory.4,22 This theory postulates that there are two distinct mental codes that result in a mental representation of any task: a verbal code and a non-verbal or imagery code. It is theorized that the verbal code is composed predominantly of three sensory modalities: vision in the form of language, auditory in the form of speech, and touch in the form of Braille.23 In comparison, the non-verbal mental code is composed of all five sensory modalities: visual,

Table 1  Baseline subject characteristics. Ordinal data analysed by Mann–Whitney test and are presented as median (IQR [range]). Nominal data are presented as percentage and analysed by Fisher’s exact test

<table>
<thead>
<tr>
<th></th>
<th>Mental practice group (n=19)</th>
<th>Control group (n=21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (%)</td>
<td>63.2</td>
<td>57.1</td>
</tr>
<tr>
<td>Previous simulations</td>
<td>6 (4–7 [1–8])</td>
<td>4 (3–5 [1–7])</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>29 (28–31 [26–40])</td>
<td>29 (28–31 [25–40])</td>
</tr>
<tr>
<td>PGY level</td>
<td>3 (2–4 [1–5])</td>
<td>2 (2–3.5 [1–5])</td>
</tr>
</tbody>
</table>

Fig 2 Box plot demonstrating the Ottawa GRS score (minimum score 0, maximum score 35) for the control and intervention groups. The central bar represents the median, the box the inter-quartile range, and the whiskers the range. Scores for each subject are the mean of the scores assigned by the two raters.

Fig 3 Box plot of modified mMIQ score (minimum score 8, maximum score 56) for the control and intervention groups. The central bar represents the median, the box the inter-quartile range, and the whiskers the range.
objects, environmental sounds, the touch of objects, and smell and taste memories. MP scripts for technical skills include discrete visual cues such as visualizing instruments, tactile cues such as moving towards an object, and cognitive cues such as distinguishing arteries from veins. As such, warm-up with MP for a technical skill may activate the verbal code by having the subject read and rehearse the verbal script. The non-verbal code may thus be activated by having subjects visualize the steps of the task in sequence, imagine how a movement may feel, or visualize how an anatomical structure may appear. It has been demonstrated that MP, when using both verbal and non-verbal cognition, results in improved performance of technical skills. These technical tasks have both a large physical and cognitive component. However, although CRM skills require a large cognitive load, the physical technical requirements are minimal. A meta-analysis of MP demonstrated that both cognitive and physical tasks improved with MP; however, the types of tasks analysed were not truly equivalent to CRM skills. For example, the cognitive tasks studied included dart throwing which still clearly requires a physical component to be successful. One explanation for a lack of difference between the groups in our study may be that non-technical skills cause activation of the verbal code due to their large cognitive component and verbal rehearsal; they may not contain sufficient stimuli from the other senses to sufficiently activate the non-verbal code, which seems to be an important aspect of MP for procedural skills.

Another potential explanation for our results is the lack of task familiarity. In the field of sport psychology, where mental imagery techniques are well entrenched, it has been found that highly skilled performers benefit more from MP than those with less skill. With regards to our findings, it is possible that the MP script, based on the practice of consultants who have lots of experience in CRM, was at too high a level for some anaesthesia trainees and subsequently individuals could not form the appropriate mental images. Moreover, perhaps the haemorrhagic PEA-simulated crisis was not advanced enough for the trainees and the GRS was not able detect the difference in performances as a result of MP. Our sample size and methodology does not allow subgroup analysis of the effect of MP by post-graduate year, and this may be an appropriate avenue for further research into MP in this domain of learning. Moreover, perhaps 20 min of MP is insufficient to develop CRM skills. Responses from the MIQ may support this hypothesis as there was no significant difference between groups on question 4, which asks participants how helpful they felt the previous activity was in preparing them for a crisis. Perhaps a longer duration of warm-up with MP may be necessary in order to see an improvement. However, a meta-analysis of MP for both cognitive and physical tasks demonstrated that 20.8 min of MP is the optimal duration for MP and that rehearsing for longer in fact had a negative effect on performance.

A final explanation for the absence of effect of MP as a warm-up for CRM is that participants in the control group may have independently prepared mentally for the simulation scenarios. All participants knew ahead of time that they were going to participate in a simulation session and most University of Toronto residents are at least familiar with the general principles of CRM from previous simulation sessions. However, the median Ottawa GRS scores place the residents on average as being somewhere above ‘advanced novice’ and less than ‘competent’ in CRM, leaving clear room for improvement.
This study has several potential limitations including the potential lack of generalizability to anaesthetists with higher levels of experience. Another limitation is that all subjects knew in advance that they were going to participate in a simulation session. Ideally, we would have recruited subjects just before their simulation session; however, due to logistical restrictions, this was not feasible. Finally, although the decision to allow for 20 min of MP was based in the literature, the majority of these studies looked at the effect of MP on technical skills in relative experts. Perhaps, the time required for MP for non-technical skills is longer. In support of this, there were no significant differences found between the groups on MIQ scores, which means that participating in 20 min of MP did not increase the participants’ perceived readiness to participate in a crisis.

In conclusion, CRM-based non-technical skills do not seem to improve with warm-up using MP. Future research is needed to further elucidate which domains of skill can benefit from MP effects in anaesthesiology and how to optimally utilize MP. A better understanding of how trainee experience influences the effect of MP is required to determine how such an inexpensive tool can be best implemented.

Supplementary material
Supplementary material is available at British Journal of Anaesthesia online.

Declaration of interest
None declared.

Funding
The research was supported by the University of Toronto Department of Anesthesia Merit Award and the St Michael’s Hospital Department of Anesthesia.

References
7 Bathalon S, Dorion D, Darveau S, Martin M. Cognitive skills analysis, kinesiology, and mental imagery in the acquisition of surgical skills. J Otalaryngol 2005; 34: 328–32

Handling editor: L. Colvin