Classroom and simulation team training: a randomized controlled trial

ROBYN CLAY-WILLIAMS1, CATHERINE A. MCINTOSH2, ROSS KERRIDGE3 AND JEFFREY BRAITHWAITE4,5,6

1Faculty of Medicine, Centre for Clinical Governance Research, Australian Institute of Health Innovation, University of New South Wales, Sydney, NSW 2052, Australia, 2Hunter New England Skills & Simulation Centre, John Hunter Hospital, New Lambton Heights, NSW 2305, Australia, 3Perioperative Service, John Hunter Hospital, New Lambton Heights, NSW 2305, Australia, 4Centre for Clinical Governance Research, Sydney NSW 2052, Australia, 5Australian Institute of Health Innovation, Sydney NSW 2052, Australia and 6Faculty of Medicine, University of New South Wales, Sydney NSW 2052, Australia

Address reprint requests to: Robyn Clay-Williams, Faculty of Medicine, Centre for Clinical Governance Research, Australian Institute of Health Innovation, University of New South Wales, Sydney NSW 2052, Australia. Tel: +61-2-9385-8911; Fax: +61-2-9385-8692; E-mail: r.clay-williams@unsw.edu.au

Accepted for publication 20 February 2013

Abstract

Objective. To test the hypotheses that classroom and simulation-based crew resource management (CRM) training interventions improve teamwork attitudes and behaviours of participants and that classroom training combined with simulation-based training provide synergistic improvements.

Design. A randomized controlled trial.

Setting. Area Health Service in New South Wales, Australia.

Participants. A total of 157 doctors, nurses and midwives randomized into one of four groups, consisting of three intervention groups and a control group.

Intervention. One-day CRM-based classroom course; one-day CRM style simulation-based training or classroom training followed by simulation-based training.

Main Outcome Measures. Pre- and post-test quantitative participant teamwork attitudes, and post-test quantitative trainee reactions, knowledge and behaviour.

Results. Ninety-four doctors, nurses and midwives completed pre-intervention attitude questionnaires and 60 clinicians completed post-intervention assessments. No positive changes in teamwork attitudes were found associated with classroom or simulation training. Positive changes were found in knowledge (mean difference 1.50, 95% confidence interval (CI) 0.58–2.43, \( P = 0.002 \)), self-assessed teamwork behaviour (mean difference 2.69, 95% CI 0.90–6.13, \( P = 0.009 \)) and independently observed teamwork behaviour (mean difference 2.30, 95% CI 0.30–4.30, \( P = 0.027 \)) when classroom-only trained group was compared with control; however, these changes were not found in the group that received classroom followed by simulation training.

Conclusions. Classroom-based training alone resulted in improvements in participant knowledge and observed teamwork behaviour. The study found no additional impact of simulation training.

Keywords: crew resource management, teamwork, competency-based education, needs assessment

Introduction

Evidence suggests that aviation-style crew resource management (CRM) teamwork training is transferable to health care [1–9]. This type of training is often recommended as a way to improve patient safety [10, 11]. Several classroom and simulation-based team training programmes, adapted from aviation CRM, have been implemented in health care settings in the USA, the UK, Europe and Australia—some in specific specialties such as anaesthesia [7, 9, 12], some in multidisciplinary settings [1, 5, 13–15] and some in the military [16, 17]. Despite the emergence of large multicentre programmes such as TeamSTEPPS [17], research on evaluating the results of training is at a formative stage [2].

Only one study evaluating classroom and simulation training in combination was found [15], conducted in the USA,
and that study was insufficiently powered to enable conclusions to be drawn. There may be synergies from a training programme that began with classroom training and progressed to fully immersive simulation, where team skills and clinical skills could be synthesized.

**Purpose of the study**

We aimed to test the effectiveness of classroom- and simulation-based CRM training, alone and/or in combination, in improving the teamwork attitudes and behaviours of health care professionals.

**Methods**

**Research hypotheses**

It was hypothesized (Table 1) that the teamwork attitudes and behaviours of participants would improve with training and that the behaviour improvement from combining the training would be synergistic rather than additive.

**Selection of participants**

The reference population chosen was health professionals in acute hospital settings (consultants, registrars, residents, registered and enrolled nurses; working in operating theatres (OTs), intensive care units, emergency departments (EDs) and labour and delivery). The source population consisted of health care professionals within the Division of Surgery, Anaesthesia and Intensive Care, and Pain Management, the Department of Obstetrics and Gynaecology, and the ED at five hospitals within an area health service. The hospitals consisted of a 640-bed tertiary referral hospital, a 187-bed affiliated hospital, two rural referral hospitals (342 beds total) and a 75-bed district hospital.

**Determination of sample size**

Sample size was estimated using a statistical program for independent groups with a two-tailed t-test [18]. A pool of 160 health-care professionals would be needed for randomization into four groups of 40, assuming a 20% withdrawal rate, to reach statistical significance.

**Randomization**

Participants were randomly assigned to one of four equivalent-sized groups:

(i) Group A—the control group, with no participation in training,
(ii) Group B—classroom training only,
(iii) Group C—simulation training only and
(iv) Group D—classroom training, followed by simulation training.

Stratified random sampling was used to allocate participants to training groups. There was no pre-randomization—everyone from the source population who volunteered to participate was accepted into the programme. To enable representative work teams for behaviour assessment, the participants were stratified into one of six groups, according to profession and specialization: OT doctors, OT nurses, non-OT doctors, non-OT nurses, ED doctors and ED nurses. They were listed under these categories in the order in which their nominations to attend training were received. Systematic sampling was then used to randomly assign participants to one of the four training groups. Random.org [19], an established web-based random number service, generated an integer sequence for each of the six lists. Names allocated the Integer 1 were assigned to Group A, Integer 2 to Group B, Integer 3 to Group C and Integer 4 to Group D.

**Methods**

**Design, sample and data-gathering strategy.** The study used a $2 \times 2$ factorial randomized experimental design (Table 2). The research consisted of pre- and post-intervention attitude measurements, and post-intervention reaction, knowledge and behavioural measurements, mirroring similar trial designs [2, 17, 20–24]. Depending on the training group, the intervention consisted of one day of classroom-based

<table>
<thead>
<tr>
<th>CRM classroom-based training</th>
<th>CRM simulation-based training</th>
</tr>
</thead>
<tbody>
<tr>
<td>No training</td>
<td>No training</td>
</tr>
<tr>
<td>Training</td>
<td>Training</td>
</tr>
<tr>
<td>• No attitude change</td>
<td>• Positive attitude change</td>
</tr>
<tr>
<td>• Provides behaviour baseline</td>
<td>• Positive behaviour compared with no training</td>
</tr>
<tr>
<td>• Positive attitude change</td>
<td>• Positive attitude change</td>
</tr>
<tr>
<td>• Positive behaviour compared with no training</td>
<td>• Positive behaviour compared with no training or single training (combined change synergistic, not additive)</td>
</tr>
</tbody>
</table>
training, one day of simulation-based training or both. Baseline measurements were taken to record attitudes of individuals, and teamwork and safety climate of work teams. Classroom and simulation-based courses, developed from aviation CRM and health care crisis resource management [7] training, respectively, were delivered to the experimental groups.

Kirkpatrick’s [25] training evaluation framework was used to assess the efficacy of the training interventions. Data were gathered at the first three of Kirkpatrick’s four steps (reaction, learning and behaviour). Evaluation at Step 4 (results) was beyond the scope of this study. Reaction data were collected via a questionnaire, learning data were gathered via a pen-and-paper multiple-choice knowledge test and attitude assessment questionnaire and behaviour data were derived from a self-assessment and post hoc observation of team performance in video-recorded simulated clinical scenarios. Training and assessment for all the groups was completed within a 12-month timeframe.

**Intervention.** Military aviation CRM knowledge, skills and attitudes were translated into learning outcomes for health care, based on a combination of a pre-training needs analysis, a review of evidence for team training and an expert panel review [26]. The competencies were incorporated into a one-day classroom-based CRM course involving facilitated discussion, case studies, video vignettes, role-plays and practice of the teamwork skills and techniques. An existing, standardized one-day CRM-style inter-professional team training course formed the basis for the simulation training. This course addressed the same competencies as the classroom training. Descriptions of similar courses can be found in the literature [7]: the training consisted of classroom experiential work, followed by participating in, and observing, scenarios in the patient simulation facility. Participants were trained in inter-professional teams in groups of up to 10 per course. Where participants completed both one-day training courses (Group D), there was an interval of at least 2 weeks between classroom and simulation training.

**Attitude assessment.** Sexton et al’s [27] safety attitudes questionnaire (SAQ) was selected as the basis for the attitude assessment instrument. It has been shown to have satisfactory psychometric properties and be valid and reliable [27, 28]. A body of published benchmarking data for the questionnaire is available [27, 29]. Fourteen questions were added to the instrument to evaluate the attitude competency components taught. Twenty questions on organizational attitudes that were unlikely to be influenced by the training were removed. The final modified SAQ was piloted.

**Behaviour assessment.** Participant teamwork behaviours were measured via both participant self-assessment and independent observation of team performance in two standardized simulated scenarios using the Mayo High Performance Teamwork Scale (MHPTS) [30]. The MHPTS has been validated for self-assessment by novice participants [30], and aligned with CRM competencies forming this study’s learning objectives. The simulations were video taped, and transferred to DVDs in randomized order within groups, resulting in 28 scenarios of ~20 min duration each. The scenarios were analysed post hoc by six independent observers who completed one day of teamwork behaviour rater training using the MHPTS and were blinded to the intervention group. Measuring behaviour in the simulated work environment provided an experience analogous to the workplace, while allowing standardization of the test scenarios and ensuring that the behavioural competencies taught in the training were assessed. Scenarios were designed to consistently assess behavioural skills [31] in different clinical groups by slightly varying clinical details between scenarios, such as patient age and setting, while keeping teamwork objectives and scenario difficulty identical. The behavioural objectives of the first scenario were to assess the performance of the team in transferring critical information, in maintaining situational awareness and dealing with distractions and in asserting themselves effectively when challenged by another health-care professional (actor). The behavioural objectives of the second scenario were to assess team decision-making and conflict resolution behaviours when resolving an ethical dilemma involving the decision to continue or cease resuscitation.

Participants were assessed in inter-disciplinary teams of three to six members, with each team performing the two scenarios back-to-back. To minimize the impact of any previous simulation experience on the outcome, each team was provided with a standard orientation to the simulator and the simulated environment, followed by a practice resuscitation scenario, immediately prior to the assessment.

**Analysis.** The six-point Likert reactions data were entered into an Excel spreadsheet, and means and standard deviations (SD) were calculated for each of the 15 areas assessed on the evaluation.

Knowledge quiz questions were marked, allocating a ‘1’ for a correct answer, a ‘0’ where more than one response was circled for a question and a ‘0’ for an incorrect or missing answer. Question 20 was found to be ambiguous and was dropped from the analysis. Data for the remaining 19 questions were entered into an SPSS data sheet for analysis. The scores for each individual were summed, and a two-tailed t-test conducted to compare intervention groups with the control group.

Attitude data were reduced in accordance with the published guidelines [32]. The five-point Likert data were entered into an SPSS data sheet and cleaned: scores for reversely worded questions were inverted to match the valence of the remaining statements so that a favourable response would always be allocated a higher numerical value, and missing or

<table>
<thead>
<tr>
<th>CRM simulation-based training</th>
<th>CRM classroom-based training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>No training</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No training</th>
<th>Group A (control)</th>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>Group C</td>
<td>Group D</td>
</tr>
</tbody>
</table>

*Classroom training was followed by simulation training for Group D.

---

**Table 2** Study design
ambiguous responses were categorized as ‘missing data’. Statistical two-tailed t-tests were performed on responses to individual questions to compare pre-test/post-test attitudes of each group (Groups A, B, C and D) and to compare post-test attitudes of each intervention group (Groups B, C and D) with the control group (Group A). Teamwork climate, safety climate and stress recognition domain [27] data were calculated and statistical two-tailed t-tests were performed to compare pre-test/post-test attitudes of each group (Groups A, B, C and D) and to compare post-test attitudes of each intervention group (Groups B, C and D) with the control group (Group A).

Behaviour data were reduced in accordance with published guidelines [30]. The data were entered into an SPSS data sheet and cleaned: missing values for Questions 1–8 were allocated a ‘0’, and missing values for Questions 9–16 were allocated a ‘1’. For self-assessed data, participant’s ratings of the 16 behaviour statements for each scenario were summed to give a participant team score for each scenario. The mean score of the participants who comprised each team was then determined. For observed data, assessor ratings of the 16 behaviour statements for each scenario were summed, to give each team a total score for each scenario. The mean of the six assessor scores for each team scenario was then determined. Statistical two-tailed t-tests were then performed using (i) the observer means and (ii) the self-assessed team scores, to compare the behaviour of each intervention group (Groups B, C and D) with that of the control group (Group A).

All statistical analyses were performed at 5% level of significance. SPSS (GradPack 17.0 Release 17.0.0 and PASWStatistics 18.0) and Microsoft Office Excel (2007 and 2008 for Mac Version 12.2.9) programs were used for data entry and analysis.

Results

Participants

Of the 157 participants, 75 (48%) were lost to the project prior to commencement of the final behaviour assessment (see Fig. 1). Of the 75 lost participants, 32 (20%) formally withdrew from the project, largely due to resignation from the health service or schedule conflicts. The remaining 43 volunteers either did not return contact or were not able to attend any of the scheduled training courses within the 12-month timeframe of the study. Of those who participated, 51% were doctors, 46% nurses and 3% did not state their profession. Average clinical experience was 10.3 years (range <1–37 years). Reactions to the classroom training were universally positive (mean affective reactions for the modules ranged from 5.3 (SD 0.66) to 5.5 (SD 0.55) and mean utility reactions ranged from 5.3 (SD 0.65) to 5.7 (SD 0.44)); reaction data were not gathered for the

Figure 1 CONSORT 2010 flow diagram.
simulation training, as this aspect of the training had been previ-
ously evaluated with positive results.

**Knowledge**

A statistically significant difference was found in knowledge test scores between participants in the classroom-only group and the control group (mean difference 1.50, 95% CI 0.58–2.43, \( P = 0.002 \)) (see Fig. 2). Statistically significant differences were also found in the responses to individual questions requiring learners to select the best communication tool for a situation where information could be misheard or mis-
interpreted, the most appropriate response to a challenge and the best strategies to maintain situational awareness.

**Attitudes**

No positive attitude changes associated with training were found. In fact, a significant negative difference was found in teamwork climate domain in the classroom-only group when compared with the control group (mean difference 15.60, 95% CI 6.18–25.02, \( P = 0.002 \)). A pre-test/post-test compa-

rison for the group that received combined training found a statistically significant positive change in safety climate (mean difference 12.30, 95% CI 2.43–22.17, \( P = 0.016 \)) when compared with the control group (see Fig. 3); however, the 9 (41%) learners from this group who withdrew from the study following the intervention may have influenced these results.

**Behaviours**

The blinded independent observers found the teamwork behaviours of the group that completed classroom-only train-
ing to be significantly improved over those of the control group (mean difference 2.30, 95% CI 0.30–4.30, \( P = 0.027 \)). When compared with the control group, improvement was not found in the simulation only group or in the group that received combined training. Intra-class correlation (ICC) was calculated to determine agreement between raters; assump-
tions were that the same raters rated all cases and the mean of all ratings was the unit of analysis. Concordance was moderate, and there was no significant difference between raters (ICC 0.53, \( F = 14.7 \)).

Like the independent observer data, the self-assessed teamwork behavour of the group that completed classroom-
only training was significantly improved over that of the control group (mean difference 2.69, 95% to 0.90–6.13, \( P = 0.009 \)), whereas significant improvement was not found in the other groups. As found in the MHTPS validation study [30], there was also a tendency in our study for self-raters to inflate their scores.

**Discussion**

**Key findings**

This is one of the few randomized controlled studies [33] to investi-
gate health-care CRM training. Our study provides evidence that collaborative learning in a classroom environment improves the demonstrated team skills of health-care workers. These results are relevant to doctors and nurses working in ad hoc teams in acute care settings. As classroom-
based team training is less resource intensive than simulation-
based team training, this is a notable finding.

The positive utility and affective reactions to the classroom training were characteristic of the reactions found in this type of training in previous studies. As expected, the classroom group fared better than the control group in the knowledge assessment, as the test was specifically designed to measure knowledge competency components taught on the classroom course. It is likely that the classroom-specific knowledge became a little more dispersed with the addition of the simulation training. Although the competencies were the same, al-
ternate terminology was used for some concepts, and emphasis was placed on synthesizing knowledge rather than step-by-step recall.
In regard to attitudes, the finding that the classroom-trained group had a significantly worse team climate domain following the training when compared with the control group was an unexpected result and inconsistent with previous studies. Discussion with classroom-trained participants indicated that some were more critical of their workplace following the training. The potential for learning to result in a harsher self-assessment of the workplace raises doubts as to the utility of this instrument (in its current form) in evaluating this type of training. The inconclusive results for the attitude assessment, despite adequate sample size to detect any changes, also suggests that this instrument may not be suitable for evaluating attitude change in learners when only a small proportion of their co-workers also participate in the training.

Behavioural improvement in the classroom group in comparison with the control group following the training, but not in the other trained groups, was also unexpected. However, like the knowledge and attitude instruments, the MHPTS was specifically developed to measure CRM behaviours. It is possible that the compelling nature of the simulation training, where the team training is more integrated with training of technical skills, resulted in a learner who was less specifically skilled in CRM teamwork techniques.

An important feature of this study is the in vitro evaluation of behaviour using independent blinded observers. The ability to standardize scenarios afforded by this method provides a powerful tool for comparing performance between groups. Much behavioural change, particular in regard to teamwork training, has been studied in vivo, but is self-reported. Self-rated behaviour assessments have been found to not reliably assess team performance when compared with externally generated measures [34, 35]. Nevertheless, more information is required on which learners go on to adopt CRM techniques in the workplace, which techniques they adopt and why.

Lessons learnt

Valuable lessons were learnt for future team training research and programme delivery. In developing questions for educational research, consideration should be given to investigating why the training worked, rather than simply addressing ‘did it work’. Training that works in one context may not work in another if we do not understand the conditions under which it is successful [36]. Regehr [37] describes this as realigning educational research from an ‘imperative of proof’ to an ‘imperative of understanding’, thus acknowledging the critical importance of context.

Where attitudes and behaviours are used as measurands, variability of the data can be minimized by training everyone in a work area at the same time. For participants such as those in this study, who come from workplaces where only a small proportion of clinicians are trained in teamwork, measuring conceptual knowledge, behaviour (practical, synthesized knowledge) and job performance (transfer of synthesized knowledge to the workplace) are likely to provide better indications of the effectiveness of a training intervention than measuring attitudes. In addition, our data suggest that standardization of teamwork training techniques and terminology across an organization may be an important factor in the effectiveness of the training for ad hoc teams.

Strengths and limitations

Our study builds on the small body of high-quality evidence in relation to the efficacy of classroom CRM training in health care. With the increased prevalence of this type of training, it is important to ensure that its implementation is evidence-based.

While the randomised controlled trial design structure helped in overcoming pitfalls that normally limit the ability to generalize results, it had limitations. The high numbers of participant withdrawals potentially resulted in positive selection bias. The impact of the withdrawals on the statistical power of the study potentially limited the ability to identify changes in attitudes and behaviour associated with the training interventions. Selection bias may limit the generalization of findings.

Authors’ contributions

R.C.W. contributed to the literature search, study design, data analysis, data interpretation and writing. C.M. contributed to the study design, data collection, data interpretation and writing. R.K. contributed to the study design, data interpretation, study supervision and review of the manuscript. J.B. contributed to the study design, study supervision, study funding and review of the manuscript.

Supplementary material

Supplementary material is available at *INTQHC* online.

Acknowledgements

We acknowledge the valuable assistance in the training and its associated administration provided by the faculty and staff of the Hunter New England Skills and Simulation Centre.

Funding

This research was partially supported under a University of New South Wales scholarship and by National Health and Medical Research Council (NHMRC) Programme grant 568612. The funding sources did not have any role in the conduct or reporting of the study. Ethics approval for the research was granted by the Hunter New England Area Health Service Human Research Ethics Committee (HNEHREC reference numbers: 08/HNE/262 and 08/08/20/5.19) and the University of New South Wales Human Research Ethics Committee (HREC reference number: 08274).
Conflict of Interest statement

None of the authors had conflict of interest in relation to the conduct or reporting of this study.

References

32. Sexton JB, Thomas EJ. The Safety Climate Survey: psychometric and benchmarking properties. Technical Report 03–03 The...
University of Texas Center of Excellence for Patient Safety Research and Practice.


