
Editorials

Anaesthetic simulation and lessons to be learned from aviation

R.L. Helmreich MD PhD,
J.M. Davies MSc MD FRCPC*

SIMULATORS are becoming increasingly important in anaesthesia. Currently they can be classified according to which component(s) of anaesthetic competence are to be acquired. Computer screen-based devices are best suited to acquisition of knowledge. For example, the *Anesthesia Simulator Consultant* has programmes for management of critical incidents during general, regional, obstetric and paediatric anaesthesia, as well as cardiac life-support. *Body* simulation has complex models of pharmacology and physiology.¹ Part-task trainers permit trainees to acquire both knowledge and skills. For example, the *PatSim* simulator and the *Access* system both provide hands-on opportunities using standard, but limited, operating room and anaesthetic equipment to manage common problems (e.g., laryngospasm, bronchospasm).^{1,2} (Virtual reality simulators provide similar, although qualitatively different, experiences.) Medium to high-fidelity devices provide the most comprehensive possibilities – allowing acquisition of knowledge, skills and attitudes.^{1,3-5} For example, the *Anesthesia Crisis Resource Management* programme requires the anaesthetist to manage such problems as breathing circuit disconnection in real time, while dealing with other operating room (OR) personnel.⁶ The *Team Oriented Medical Simulation* system allows all OR team members to work together when situations such as excessive surgical bleeding occur.⁵

In addition, simulators are also used, to a minimal extent, for testing.⁴ This is an area which is still considered “difficult at present.”^{1,7} Part of the difficulty comes from the fact that anaesthetic simulators, for the most part, do not fully ‘simulate’ the real-world. The ‘patient’ lacks certain characteristics (although many of these, such as skin colour and temperature,

are under active development). In addition, unlike aviation’s use of standard operating procedures (SOPs) for emergencies, many anaesthetists are reluctant to accept specific management protocols for certain procedures. Part of this reluctance may stem from the fact that in anaesthesia there is little evidence that practising in one way, rather than another, may change the (measurable) outcome for the patient. One must think of the concept of ‘equifinality’ – the notion that you can get from Point A to Point B in more than one way – when an individual or a team tackles a complex problem. As has been shown in aviation, equifinality does not discount the importance of algorithms in defining and improving processes for simple problems. However, in complex situations involving multiple challenges, many solutions are possible that will lead to equivalent outcomes.

Standardising the response to specific single problems is the subject of the article in this issue by researchers from the simulator group at Sunnybrook Hospital.⁸ Because of what Devitt and coworkers described as the subjective nature of assessment of performance of anaesthetists, their goal was to develop a more objective evaluation. They developed two, one-hour scripted scenarios of five anaesthetic problems requiring diagnosis and management. These problems originated in the “anaesthetic machine,” “mesenteric traction,” “respiratory” system, “cardiovascular” system, and in “long-term monitoring.” The manifestation and consequences of each problem could be reproduced using the simulator’s computer. Three defined responses to each problem were also developed: “no response,” “compensating intervention” and “corrective treatment.” Thus, a total of 30 problems/actions were presented. Videotape recordings of

From the Department of Psychology, University of Texas, Austin, Texas 78712 and the *Department of Anaesthesia, Foothills Medical Centre, University of Calgary, 1403 29th Street NW, Calgary, Alberta, Canada T2N 2T9.

Address correspondence to: Dr. J.M. Davies.

each one hour session were made with anaesthetists who were instructed to 'role-play' in responding to each problem, which were presented randomly to each role-playing anaesthetist.

The videotapes were then reviewed by two anaesthetists who had been in active clinical practice for more than five years. Neither had been involved in the developmental part of the study although both had participated in a three day course about the general principles of simulator scenario design and construction. Both raters had also received training in the content and scoring system of this particular study. The two raters were asked to score each of the problems/actions. These scores were then compared and shown to have excellent inter-rater reliability.

We wish to comment on two aspects of this paper: content and evaluation *vs* assessment. In doing so, we will draw from lessons learned in aviation simulation, on which most anaesthetic and operating room simulation is based.

Content

Devitt and coworkers based their study on two scenarios which they had developed. These scenarios consisted of five problems, not described apart from the topic, but which were considered "clinically relevant."⁸ While these problems might be appropriate for the goal of this particular study, 'testing the testers,' there may be difficulties with the use of these apparently unrelated problems in on-going use in simulator-based instruction. Most critical incidents and accidents in anaesthesia, as in aviation, do not result from five separate crises, but rather from the compounding of simple errors, what Gaba has termed the "chain of accident evolution."⁹

In aviation, simulator training includes 'Line Oriented Flight Training' or LOFT – where sessions take the form of real-time, normal 'line' flight. "High-technology, real-life simulation" is used to "duplicate the operational environment for flight crew training purposes."¹⁰ Problems encountered might include "passenger illness, complex weather problems or sudden unavailability of destination airport and a resulting fuel problem."¹¹ For example, in one LOFT session, before take-off the pilots learn that fog ahead of a cold front is causing low visibility at both Destination #1 and the 'alternate' destination. In addition, a piece of equipment, an auxiliary power unit (APU), will not start until a mechanic starts it. Take-off and cruise are then uneventful. On the initial approach to Destination #1, a generator warning light comes on. The crew have now finished their approach check-list but must continue to monitor the warning light. The weather now deteriorates, landing is not possible and a missed

approach manoeuvre is undertaken. Because another aircraft is in the only 'published' holding position, the crew must fly their plane into a 'non-published' holding position. The generator then fails as the crew enters this holding position. The APU then will not restart and the crew must evaluate the severity of the electrical problem and make contingency plans, including whether or not to continue to hold or to seek another destination.¹⁰ Thus, they face a complex situation with high work-load – a time when communication, decision-making, problem-solving and conflict resolution are most critical. This scenario, therefore, consists of one mechanical problem and a weather-related problem, both linked in the context of the flight.

Devitt *et al.* do not state what the ultimate use of these scenarios is to be. One possible use is for the simulation of any one of the five separate problems. However, using the model set by aviation, to present a scenario with five separate problems would probably be considered unrealistic.

We would also question what the "corrective treatment" represents, as opposed to the "compensating intervention" or even the "incorrect action."⁸ We would urge caution at making the "corrective treatment" the gold standard of crisis management, particularly when the authors use this term interchangeably with "Definitive Management." In a previous editorial, we commented that human factors programmes, (which includes anaesthetic simulation), need to be "data driven."¹² We must therefore ask: are the salvage protocols (crisis management) actually based on data? How can the authors reassure us that these are indeed the right things to do and not just "clinically reasonable" without some form of validation, other than that described in this article?

Evaluation versus assessment

The authors base their work on the premise that they wished to have a more objective means of evaluating anaesthetists.⁸ It is important to distinguish assessment from evaluation, as has already been addressed in aviation, particularly in the area of Crew Resource Management (CRM).¹³ In aviation, assessment of CRM provides information about team responses to complex situations with respect to both technical and human factors. Technical factors represent the manual 'stick and rudder' skills necessary to fly the aircraft. Human factors represent interpersonal behaviours, such as situation awareness, decision making and team coordination. All information obtained about individuals is considered confidential and 'without jeopardy'. Data provide an understanding of both strengths and weaknesses, which, over the course of several assessments, can be used to determine trends.

In contrast, evaluation is the formal process of certification and continuing qualification. Evaluation is the last stage of a process which needs to be introduced incrementally: first teach, then assess, then evaluate. It should be noted that formal evaluation of team performance in aviation operational simulations, (known as Line Operational Evaluation or LOE), is not undertaken until crews have gained experience and have received assessment and feed-back from the non-jeopardy LOFT training. In addition, LOFT is not undertaken until crews have mastered basic technical skills.

Indeed, as mentioned above, the anaesthetic culture may not be ready for evaluation. However, previous work from the Sunnybrook group would suggest that Canadian anaesthetists are very supportive of the use of simulators for training. The same study showed that there was much less support for "compulsory use for recertification," although residents were more accepting of this concept than were those who had completed their training.¹⁴

In this regard, evaluation of individual decision-making by anaesthetists is more reasonable, particularly when one considers that the problems developed for the present study actually represent 'part-task training' – the individual mastery of basic knowledge and skills. As such, Devitt *et al.* are using a medium to high fidelity simulator to provide part-task training to individual anaesthetists. The Sunnybrook simulation programme uses as its basis the *Anesthesia Crew Resource Management* (ACRM) programme, which is stated to be based on aviation's Crew Resource management.⁶ Actually, ACRM is a programme to assess the ability of an individual 'in the hot seat' to handle specific anaesthetic crises, although 'resource' management and communication are an integral part of the programme. In contrast, true CRM addresses team-interface issues – those problems which most often arise between anaesthetists and surgeons – and which require an understanding of how teams make decisions.⁵ We would suggest that there is a need for both individual and CRM (team) programmes, and that the two should be clearly differentiated.

In summary, we agree that the goals of the present study are worthwhile but recommend that the content and the process of implementation be reconsidered. There is a need to provide opportunities for practice and assessment until the culture supports the fairness of the assessment process. This applies both to the individual practitioner (as with part-task training) and with the entire operating room team (as with CRM). At that point, formal evaluation, leading to a programme of certification and recertification, can be considered and implemented.

La simulation en anesthésie : les leçons à tirer de l'aviation

En anesthésie, les simulateurs prennent de plus en plus d'importance. On les classe présentement en fonction du domaine de la compétence requise. Les dispositifs avec écran sont les mieux adaptés à l'acquisition des connaissances. Par exemple l'*Anesthesia Simulator Consultant* offre des programmes orientés vers la prise en charge d'incidents critiques pendant l'anesthésie régionale, obstétricale et pédiatrique ainsi que l'initiation de la réanimation cardiaque. La simulation sur mannequin introduit des modules pharmacologiques et physiologiques complexes.¹ Des instructeurs spécialisés permettent à l'étudiant d'acquiescer à la fois connaissance et compétence. Par exemple, le simulateur *Patsim* et le système *Access* procurent tous les deux la possibilité de gérer des problèmes courants (ex. laryngospasme, bronchospasme) en manipulant dans certaines limites l'équipement en usage dans la salle d'opération et en anesthésie.^{1,2} La réalité virtuelle fournit aussi des expériences de simulation identiques quoique différentes au point de vue de la qualité. Les modules à fidélité moyenne à haute offrent les plus grandes possibilités dans le domaine de l'acquisition des connaissances, des compétences et des attitudes.^{1,3-5} Par exemple, le programme de gestion des situations de crise (*Anesthesia Crisis Resource Management*) oblige l'anesthésiste à prendre charge en temps réel de problèmes tel le débranchement du circuit, tout en communiquant avec les autres membres du personnel des salles d'opération.⁶ Le système de simulation orienté vers l'équipe médicale (*Team Oriented Medical Simulation*) favorise le travail collectif de tous les membres de l'équipe quand une situation grave comme une hémorragie massive survient.⁵

Jusqu'à maintenant, les simulateurs ont été très peu utilisés pour les examens (évaluation normative).⁴ Cet aspect a toujours été considéré comme difficilement réalisable.^{1,7} Une partie du problème provient du fait que les simulateurs anesthésiques n'imitent pas la réalité parfaitement. Le «patient» n'offre pas toutes les caractéristiques requises (bien que certaines comme la couleur de la peau et la température soient en voie de développement). En outre, contrairement à l'aéronautique qui utilise des modalités d'interventions standardisées, plusieurs anesthésistes refusent de se plier à des protocoles de prise en charge pour certaines interventions spécifiques. Une

partie de cette résistance peut provenir du fait qu'en anesthésie, il n'est pas évident que pratiquer d'une façon plutôt que d'une autre a quelque influence sur les résultats mesurables. Il faut tenir compte de «l'équifinalité» – le concept selon lequel il est possible, en face d'un problème compliqué, qu'un individu ou un groupe d'individus se rende du point A au point B en empruntant des chemins différents. Comme l'a démontré l'aviation, l'équifinalité ne renie pas l'importance des algorithmes pour la résolution et la correction de problèmes simples. Par contre, dans les situations complexes créées par l'association de plusieurs problèmes, il est possible que plusieurs solutions aboutissent à des résultats équivalents.

Ce mois-ci, dans le *Journal*, le groupe de chercheurs en simulation de l'hôpital Sunnybrook publie un travail sur l'uniformisation des réactions à des problèmes spécifiques.⁸ À partir de leur description de la nature subjective de l'évaluation de la performance des anesthésistes, Devitt *et al.* se sont donnés pour objectif d'élaborer un mode d'évaluation plus objectif. Ils ont développé deux scénarios d'une heure sur des problèmes anesthésiques exigeant un diagnostic et sa prise en charge. Les problèmes proposés mettaient en cause l'appareil d'anesthésie, les tractions mésentériques, le système respiratoire, le système cardiovasculaire et le monitoring à long terme. Les manifestations et les conséquences de chaque problème étaient transmises par simulation informatisée. Trois réactions possibles avaient été développées : «aucune», «intervention neutralisante» et «correction thérapeutique». Ainsi, un total de 30 problèmes/réactions étaient disponibles. Au cours de sessions d'une heure, on enregistrerait sur vidéocassette les réactions des anesthésistes «en charge» mis en présence de chacun des problèmes choisis au hasard.

Les vidéocassettes étaient ensuite révisées par deux anesthésistes possédant cinq années d'expérience et plus. Aucun des deux n'avait participé à l'élaboration de l'étude bien qu'ils aient tous les deux suivi un cours de trois jours sur les principes généraux de la création et de la construction des scénarios informatisés. Les deux évaluateurs avaient reçu une formation spéciale sur le contenu de l'étude et le système de notation.

Nous désirons commenter deux aspects de cet article : le contenu et l'évaluation cognitive vs l'évaluation normative. Pour y arriver, nous allons faire appel aux leçons de la simulation en aviation sur laquelle se base la majeure partie de la simulation en anesthésie et en salle d'opération.

Le contenu

Devitt et ses collègues ont basé leur étude sur deux scénarios développés par eux-mêmes. Ces scénarios comprenaient cinq problèmes de portée limitée qu'ils

jugeaient applicables à la clinique.⁸ Alors que ces problèmes peuvent concorder avec les objectifs de cette étude qui était l'évaluation des évaluateurs, il peut être difficile d'accorder ces problèmes sans relation apparente entre eux avec les besoins de l'enseignement basé sur la simulation. Les plupart des accidents critiques, comme en aviation, ne résultent pas de cinq crises séparées mais plutôt de la combinaison d'erreurs simples appelée par Gaba «chaîne de l'évolution de l'accident».⁹

En aviation, l'apprentissage par simulateur comprend des sessions de vol de ligne en temps réel (*Line Oriented Flight Training* ou LOFT). Une simulation réaliste de haute technologie sert à reproduire un environnement opérationnel pour la formation du personnel volant. Les problèmes rencontrés peuvent comprendre un passager malade, des conditions atmosphériques complexes et un épuisement consécutif du carburant.¹¹ Par exemple, au cours d'une session de LOFT, avant le décollage, les pilotes sont avertis qu'un front froid provoque une baisse de visibilité à la fois au point de destination N°1 et à son alternative. De plus, une pièce d'équipement, une génératrice auxiliaire ne démarrera pas sans l'aide d'un mécanicien. À la destination N°1, pendant la manoeuvre d'approche préliminaire, le signal d'alarme de la génératrice s'allume. L'équipage a terminé sa liste de vérification tout en gardant un oeil sur le voyant d'alarme lumineux. Le climat se détériore, l'atterrissage est impossible et il faut avorter la manoeuvre d'approche. Parce qu'un autre avion se trouve dans le corridor d'attente affiché, l'équipage doit conduire l'appareil dans un autre corridor non affiché. Au moment où l'appareil pénètre dans cet espace, la génératrice flanche. La génératrice auxiliaire ne démarre pas et l'équipage doit juger de la gravité du problème électrique et trouver une solution de rechange, y compris la prolongation de l'attente ou la recherche d'une autre destination.¹⁰ L'équipage fait donc face à une situation très complexe exigeant un surcroît de travail – au moment où les communications, les prises de décision et la résolution du conflit sont critiques au plus haut point. En résumé, ce scénario provoque un problème mécanique doublé d'une complication météorologique au cours d'un vol aérien.

Devitt *et al.* ne définissent pas à quoi serviront leurs scénarios. Peut-être désirent-ils simuler chacun des cinq problèmes individuellement. Cependant, dans le contexte du modèle que nous venons de décrire, il est probable que la présentation de cinq problèmes séparés serait considérée comme chimérique.

Nous pouvons aussi nous demander ce que signifie l'expression «traitement correctif» comparativement à «l'intervention compensatrice» ou même «l'action fautive».⁸ Nous recommandons la prudence lorsqu'il s'agit de faire du «traitement correctif» la règle d'or de la ges-

tion d'une crise, spécialement quand les auteurs font alterner indifféremment ce terme avec celui de «gestion définitive». Dans un éditorial antérieur, nous avons expliqué que les programmes où les facteurs humains interviennent (incluant la simulation anesthésique) doivent être activés par la reconnaissance de données.¹² Nous devons donc nous demander si les protocoles de sauvetage (gestion de crise) sont actuellement basés sur ces prémisses. Comment les auteurs peuvent-ils nous confirmer que se sont les choses qu'il faut faire et qu'elles ne sont pas uniquement raisonnables dans le sens clinique sans autre forme de validation que celle que décrit cet article ?

Évaluation normative vs cognitive

Les auteurs justifient leur travail sur les prémisses qu'il faut des critères plus objectifs pour évaluer l'anesthésiste.⁸ Il est important de distinguer entre l'évaluation cognitive et normative, comme on l'a fait dans l'aviation, particulièrement dans le domaine de la gestion des capacités de l'équipage (*Crew Resource Management* – CRM). Dans l'aviation, l'évaluation cognitive par CRM fournit des renseignements sur la réponse de l'équipage à des situations complexes en tenant compte à la fois des facteurs humains et des facteurs techniques. Les facteurs techniques sont représentés par la maîtrise du pilotage comme la manipulation des contrôles de direction et du gouvernail. Les comportements interpersonnels comme l'appréciation de la situation, la prise des décisions et la coordination entre les membres de l'équipe constituent les facteurs humains. Tous les renseignements concernant l'équipage sont considérés confidentiels et non compromettants. Les données permettent de connaître les forces et les faiblesses de chacun et, après plusieurs séances d'évaluation, serviront à déterminer des tendances.

Par contre, l'évaluation normative constitue un processus formel et continu de certification et de qualification. L'évaluation normative est le dernier stage d'un processus qui doit être introduit par étape : d'abord l'enseignement, ensuite l'évaluation des connaissances et finalement la comparaison avec des normes. Il est à noter qu'en aviation, l'évaluation normative de la performance d'un équipage dans les simulations opérationnelles connue comme l'évaluation opérationnelle de ligne (*Line Operational Evaluation* ou LOE) n'est entreprise qu'après l'acquisition de l'expérience par l'équipage, après son évaluation cognitive et après le feed-back provenant de la formation par LOFT. En outre, le LOFT n'est commencé qu'une fois l'habileté technique de base acquise.

Le monde de l'anesthésie n'est peut-être pas mûr pour l'évaluation normative par la simulation. Par contre, le travail admirable effectué par le groupe du Sunnybrook suggère que les anesthésistes canadiens appuient forte-

ment l'utilisation des simulateurs pour l'apprentissage. La même étude a montré que l'enthousiasme diminuait lorsqu'il s'agissait de la recertification obligatoire, bien que les résidents aient été beaucoup moins hostiles à ce concept que ceux qui avaient terminé leur formation.¹⁴

À cet égard, l'évaluation de la prise de décision individuelle par les anesthésistes est plus raisonnable, surtout si on considère que la présentation des problèmes élaborés pour cette étude représentent actuellement une forme d'apprentissage limité – la maîtrise individuelle des connaissances et des habiletés fondamentales. Dans ce but, Devitt *et al.* utilisent un simulateur de fidélité moyenne à haute pour procurer à des anesthésistes individuels des exercices d'apprentissage limités. Le programme de simulation de Sunnybrook a comme ressource de base le programme de gestion des ressources de l'équipe anesthésique (*Anesthesia Crew Resource Programme* – ACRM) qu'on dit lui-même calqué sur le programme de gestion de ressources de l'équipe volante (CRM).⁶ En réalité, l'ACRM est un programme qui évalue l'individu installé au poste de commande⁵ pour la prise en charge de crises anesthésiques spécifiques, bien que le domaine de la gestion des ressources et des communications fassent aussi partie intégrale du programme. Par contre, le «vrai» CRM est orienté vers des problèmes d'interfaces entre membres d'un équipage – le type de problème qui surgit souvent entre anesthésistes et chirurgiens – et qui nécessite une connaissance de la façon de prendre des décisions en équipe.⁵ Nous croyons à l'utilité de deux programmes axés sur les individus et sur les équipes (CRM) mais les deux devraient être clairement différenciés.

En résumé, nous sommes d'accord sur la valeur des objectifs de la présente étude mais nous recommandons la révision de son contenu et de son implantation. Il faut continuer d'innover pour améliorer la pratique et l'évaluation jusqu'à ce que le mécanisme d'évaluation soit jugé équitable. Ceci s'applique autant à l'individu (comme avec les exercices limités) qu'à toute l'équipe de la salle d'opération (comme avec le CRM). À ce moment-là, il sera possible de considérer et d'implanter une évaluation normative par simulation pour la certification et de la recertification.

References

- 1 Chopra V. Anaesthesia simulators. In: Aitkenhead AR (Ed.). *Quality Assurance and Risk Management in Anaesthesia*. Baillière's Clinical Anaesthesiology International Practice and Research. London: Baillière Tindal 1996; 10: 297–315.
- 2 Byrne AJ, Hilton PJ, Lunn JN. Basic simulations for anaesthetists. A pilot study of the ACCESS system. *Anaesthesia* 1994; 49: 376–81.

- 3 *Gaba DM, DeAnda A.* A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. *Anesthesiology* 1988; 69: 387-94.
- 4 *Chopra V, Engbers FHM, Geerts MJ, Filet WR, Bovill JG, Spierdijk J.* The Leiden anaesthesia simulator. *Br J Anaesth* 1994; 73: 287-92.
- 5 *Helmreich RL, Davies JM.* Human factors in the operating room: interpersonal determinants of safety, efficiency and morale. *In: Aitkenhead AR (Ed.). Quality Assurance and Risk Management in Anaesthesia. Bailliere's Clinical Anaesthesiology International Practice and Research.* London: Baillière Tindal 1996; 10: 277-95.
- 6 *Howard SK, Gaba DM, Fish KJ, Yang G, Sarnquist FH.* Anesthesia crisis resource management training: teaching anesthesiologists to handle critical incidents. *Aviat Space Environ Med* 1992; 63: 763-70.
- 7 *Gaba DM.* Human work environment and simulators. *In: Miller RD (Ed.). Anesthesia, 4th ed.* New York: Churchill Livingstone 1994: 2635-79.
- 8 *Devitt JH, Kurrek MM, Cohen MM, Fish K, Fish P, Murphy PM, Szalai J-P.* Testing the raters: inter-rater reliability during observation of simulator performance. *Can J Anaesth* 1997; 44: 924-928.
- 9 *Gaba DM, Maxwell M, DeAnda A.* Anesthetic mishaps: breaking the chain of accident evolution. *Anesthesiology* 1987; 66: 670-6.
- 10 *Maurino DE.* Human factors training in aviation. *In: Telfer RA (Ed.). Aviation Instruction and Training.* Aldershot: Ashgate Publishing Limited, 1993: 95-113.
- 11 *Mellor A.* Design, development and implementation of a CRM program. *In: Telfer RA (Ed.). Aviation Instruction and Training.* Aldershot: Ashgate Publishing Limited, 1993: 368-84.
- 12 *Davies JM, Helmreich RL.* Simulation: it's a start (Editorial). *Can J Anaesth* 1996; 43: 425-9.
- 13 *Helmreich RL, Foushee HC.* Why crew resource management? Empirical and theoretical bases of human factors training in aviation. *In: Wiener EL, Kanki BG, Helmreich RL (Eds.). Cockpit Resource Management.* San Diego: Academic Press, 1993: 3-45.
- 14 *Kurrek MM, Fish KJ.* Anaesthesia crisis resource management training: an intimidating concept, a rewarding experience. *Can J Anaesth* 1996; 43: 430-4.