The FAA Center of Excellence for Technical Training and Human Performance

AJI-2 Project #20 Benchmarking Premier Technical Training Providers – lessons from other industries

Final Report

Submitted by

Massoud Bazargan, PhD
Professor of Production Operation
Department of Management, Marketing and Operations
Email: bazargam@erau.edu
Phone: 1 386 226 - 6705

Jayendra Gokhale, PhD
Assistant Professor of Economics
Department of Accounting, Economics, Finance and Information Systems
Email: gokhalej@erau.edu
Phone: 1 386 226 4967

Chunyan Yu, PhD
Professor of Air Transport Management
Department of Management, Marketing and Operations Management
Email: yuc@erau.edu
Phone: 1 386 226 6723

With Assistance of our MBA students
Ziyi Huang,
Nidihi Trambadia
Marissa Pecora

College of Business
Embry-Riddle Aeronautical University
600 S. Clyde Morris Blvd
Daytona Beach, FL 32114

Table of Contents

Table of Figures

Executive Summary

There are several critical industries in which one mistake may mean an irreversible loss of human life. Aviation is one such key industry. Others include first responders and law enforcement officers and surgeons. It is critical that when training these professionals for job performance, all aspects of their role be considered including physical, mental and psychological. There are occasional events which lead to potential life threatening consequences, while others are business as-usual or no business at all. This report is intended to keep these roles in mind and present to the Federal Aviation Administration, some key lessons learned from the above mentioned industries and how they relate to the training of air traffic controllers (ATC) and air transportation systems specialists (ATSS).

There are primarily three areas where the industries reviewed in this report identify as crucial for their staff trainings as follows:

- Intensive selection of candidates – The candidates are driven with ability to think outside the box in order to deal with unexpected and undefined events. Some of major metrics in candidate selection include, being healthy, self-motivated, have reasonable psychomotor skills and emotional stability.
- Constant updates to training curricula and assessments– Traditional lectured-based one size-fit-all curricula are found to be inefficient. The training programs and curricula need to be revised continuously to be relevant, innovative and data driven. Similarly, the assessments need be innovative, relevant and focused towards evaluating technical skills
- Adoption of technology – Technology has proved to be very crucial in training. These are in the form or augmented and virtual reality, gamification and serious gaming, situational awareness and self-awareness to cope with exposure to sudden life threatening and potentially stressful situations.

The report presented in the following sections examine each of the above areas and other relevant successful training methods in different industries.

1. Background and Introduction

The original objective of ***AJI-2 Project #20 Benchmarking Premier Technical Training Providers*** was to develop benchmark metrics or key performance indicators (KPI) for provision of technical/professional training in general and aviation technical workforce in particular, and to apply these metrics (or KPIs) to measure and compare technical training providers and identify the best practices. Our research team spent substantial time and efforts conducting literature review on methodologies and case studies for benchmarking technical training programs, including assessment programs of air navigational service providers in other countries. The scope of the research, however, was changed at the direction of FAA sponsors in February, 2018. The revised objective of ***AJI-2 Project #20 Benchmarking Premier Technical Training Providers – lessons from other industries*** is to identify the best practices of technical training programs including tools and evaluation processes in selected industries that are applicable to ATC and ATSS technical training. Our research efforts to achieve the revised objective are based mostly on literature review, including both academic journal articles and industry practitioners' reports.

Brady and Stolzer (2017) state that over 12,000 air traffic controllers are expected to retire, resign, be promoted or lost to attrition between 2012 and 2021, which means that FAA will need to fill these positions with people "who grew up in a digital world with a preference to digital technologies". That is, training for these new professionals must keep up with the fast changing digital landscape. Therefore, our literature review focuses on the applications of non-traditional technologies in providing technical training.

This report documents our research team's efforts and presents research findings to achieve the revised objective, and does not include findings from our earlier efforts based on the original project objective[1]. The rest of the report is organized as follows: Section 2 reports on results from an expert focus group discussions, and identifies the industries that are potentially relevant to Air Traffic Controller (ATC) and Airway Transportation Systems Specialist (ATSS) training; Section 3 provides an overview of a number of non-traditional and innovative technology based training methods that have been used for technical training in various industries, as suggested by the experts

[1] The literature review on benchmarking methodologies and practices from our earlier efforts are not presented in this report in order for the report to have a more coherent structure.

from our focus group; Section 4 discusses advanced technical programs in the healthcare industry (nursing and surgery) and implications for ATC and ATSS training; Section 5 reviews the advanced practices in pilot training; Section 6 discusses technical training for first responders and emergency personnel; Section 7 reviews relevant training programs in some other industries including nuclear power plants, agriculture, military and labor markets. Section 8 offers overall recommendations and best training practices and also provides concluding remarks.

2. Focus Group Discussion and Selection of Comparable Industries

To get a better understanding of the ATC and ATSS training programs, the research team formed a focus group composed of retired FAA air traffic controllers who are now faculty members at Embry-Riddle Aeronautical University (ERAU). This is our "second best" solution, as ideally, we would like to form one or two focus groups consisting of current air traffic controllers, and instructors at FAA training facilities. Unfortunately, we were informed that access to current ATC and ATSS personnel was not possible.

Experts in the focus group include:

- Dr. William Coyne, Professor and Undergraduate Program Coordinator, Air Traffic Management
- Professor Martin Lauth, Associate Professor of Air Traffic Management
- Dr. Gregory McGuirk, Associate Professor and Chair, Department of Applied Aviation Sciences.

They are experienced former air traffic controllers, retired from FAA, and are now teaching air traffic control and air traffic management courses at ERAU. The research team briefed the experts on the team's research efforts up to that point, and they provided valuable feedback on the direction of our next step.

The focus group helped research team to define the main characteristics of the technical tasks ATC and ATSS personnel have to perform on their respective jobs, and to identify other occupations that share similar characteristics. The following summarizes the key points from the meeting with the focus group:

- Air Traffic controllers (ATC) have to carry out their daily tasks often in a fast paced environment, but may also have lengthy down times depending on their locations. There is a high risk associated with their daily tasks, and potential for the loss of lives as a consequence of the decisions they make or how they execute their tasks. Surgeons, first responders including emergency medical technicians, firefighters, and police, and pilots share many similar characteristics in their daily work.

- Airway Transportation Systems Specialists (ATSS) are responsible for maintaining the communications, navigation, surveillance, and automation equipment. They are trained to work with a wide range of equipment. Very often the ATSS sit and wait for something to happen, then their role becomes critical. Most of these equipment have backup, and when an equipment fails, it is often replaced, instead of extensive repair. Potential for loss of life only occurs when an equipment has a major mishap or failure. YouTube videos and/or training videos created by manufactures are good ways to provide training, especially refresher training. Nuclear facility technicians share many similar characteristics in their training and re-training requirement as ATSS.

- The focus group also touches on how NextGen may affect both ATC and ATSS training. With the increasing applications of various forms of artificial intelligence, the role of human intervention is beginning to diminish. Therefore, people may be trained to monitor the system, rather than actively to operate it.

Based on our discussions with the experts in our focus group as well as our precursory literature review, we decided to focus on training programs for surgeons and some other hospital personnel, first responders, and pilots. The focus group also re-confirmed our initial premise of focusing on innovative and technology based methods as training tool.

3. Innovative and Technology based Training Methods

SURFnet (2016) maps out 13 individual technological trends of innovative training and education, including virtual reality, serious gaming, and gamification. This section reviews and discusses each of these training methods.

3.1 Virtual Reality

Using a computer or smart device, Virtual Reality (VR) simulates reality in order to engage users in sensory experiences. With a VR headset, he/she is inside the 3D world, which enables the user to actually walk around and use controllers to interact in a natural way with objects in the environment.

Virtual reality has been used to train surgeons since early 2000s, and will be discussed in more detail in Section 4. It has also been used in other industries for both initial training and refresher training. Vora et al. (2002) outline the development of a virtual reality (VR) system for aircraft inspection training. Reliability of human inspectors is fundamental to an effective aircraft maintenance system, and 90% of all inspections for aircraft maintenance is visual in nature. An effective training program is thus critical. The authors find that the VR system "suitably mimics the real world environment", and is preferred by trainees over an existing 2D PC based simulator with a keyboard and a two-button mouse. They conclude that the VR system has the potential for use as an off-line training tool for aircraft visual inspection tasks.

Pinheiro et al. (2014) present the development of a 3D mechanical maintenance training simulator for F-16 engines based on the maintenance requirements of Portuguese Air Force, using a low-cost simulation platform and a software architecture that separates simulation control from simulation visualization. At the Portuguese Air Force, engine technicians go through an initial training process at the Centre for Military and Technical Training of the Air Force, and then are assigned to different air bases to receive further training, focusing on the specific engines and aircraft deployed and serviced there. The "on-the-job" training phase is resource-demanding, and at a higher risk of costly component damage during training. The research team attempted to develop a 3D multi-user mechanical training simulator on a low-cost simulation platform and a software architecture that separates simulation control from simulation visualization that provides

trainees with more opportunities on training, which allows trainees to be better prepared when they start their on-the-job training. Consequently, the simulator helps to optimize the effectiveness of the resource-intensive training occasions with physical engines.

Virtual reality (VR) has also been used as a tool for a wide range of trainings for fire fighters. There are numerous academic articles on the subject, and many commercially available VR training systems. For example, Cha et al. (2012) develop a fire training simulator that is capable of calculating various invisible physical quantities such as toxic gases and heat as well as visible factors such as smoke and flame, which allows trainees to intuitively experience dangerous fire environments. Xu et al. (2014) attempt to create a realistic and accurate smoke environment for effective virtual training, which allows trainees to assess the safety of different paths for evacuation or rescue and to identify the safest path. Section 6 provides more discussions on first responder training.

Examples of commercial VR training systems include FLAM Trainer (Australia), LUDUS (Spain), and ADMS-Fire (Orlando, FL, USA). These systems can be adapted to a particular training program's need, creating customized scenarios and risk situations. There are also numerous VR training programs specifically created for certain situations. For example, Vector Solutions partnered with Pasco County Fire Rescue to create firefighter safety training courses based on a 360 Virtual Reality video shoot. The army combat training system, VIRTSIM (Virtual Training Simulator), created by Raytheon, is an immersive, open space, VR training ground. It uses a basketball court-sized rubber game pad and a weapon-mounted controller to keep track of a soldier's movements. VIRTSIM is also used by Malaysian Army and UAE. NASA's Hybrid Reality Laboratory uses Earthlight, a VR game based on the authentic depiction of human spaceflight, in its next-generation training systems for astronauts, engineers and scientists. More discussions on VR training programs are provided in the subsequent sections.

To provide better understanding of Virtual Reality, Appendix A discusses some commonly used terms and definitions associated with virtual reality and its application in simulation based learning.

3.2 Serious Gaming

Serious games are games that have been specifically developed as learning and training tools. Serious gaming takes place in a virtual environment in which the player can perform actions and see the result of these actions directly. Gaming elements ensure that players are motivated to achieve their objective.

The armed forces have been using military games, such as ***America's Army***, for many years. ***America's Army*** is a game technology platform used to develop first-person shooter (FPS) video games published in 2002 by the U.S. Army. The game allows Americans to virtually explore the Army at their own pace and according to their interests to determine if soldiering matches their needs, interests and abilities. ***America's Army*** represents the first large-scale use of game technology by the U.S. government as a platform for strategic communication, recruitment, and the first use of game technology in support of U.S. Army recruiting. Other examples of military games include Super Nintendo ***M.A.C.S*** (The Multi-Purpose Arcade Combat Simulator), Xbox ***Full Spectrum Warrior***, and online ***Strike Group Defender: The Missile Matrix.*** M.A.C.S is a shooting game developed for the U.S. Army as a cheap way to train shooting skills. Full Spectrum Warrior is a squad-based tactics game in which the player issues commands to two fireteams, Alpha and Bravo. There are two versions of the game, an entertainment version and an Army version that is accessible through a static unlock code. ***Strike Group Defender: The Missile Matrix*** challenges trainees (sailors) to react to cruise missiles heading their ships. Trainees can choose to intercept the missiles or to deploy decoys to divert the missiles from their targets.

Video games have also been used in training first responders. For example, ***EDGE*** (Enhanced Dynamic Geo-Social Environment) is a multiplayer online training platform for first responders in a coordinated response to critical incidents. Its first scenario, an active shooter incident at a hotel, brings law enforcement, fire, emergency medical services, dispatch, and unified command together. The goal of EDGE is not to train first responders to perform their respective job duties, but to sharpen their decision-making skills when they do their jobs during a critical incident. ***ZERO HOUR: AMERICA'S MEDIC*** is window based single player game to train first responders to respond to mass casualty incidents such as earthquakes and terrorist attacks. It is a mission based game, with each mission starting with the trainee/player in an ambulance and

receiving calls from the EMS dispatcher. As the ambulance approaches the scene, the trainee sees the chaotic scene unfolding. The trainee must choose which equipment to bring and then acts in accordance with the unique requirements of the situation. ***Medrills*** is a series of mobile-based iOS and Android applications that combine 3D visualizations and gaming technology to provide an inquiry based, immersive learning environment for EMS, EMT, paramedic, and combat medic training on the go. It includes applications such as forming a scene assessment, identifying immediate life threats, and etc. More discussions can be found in Section 7.

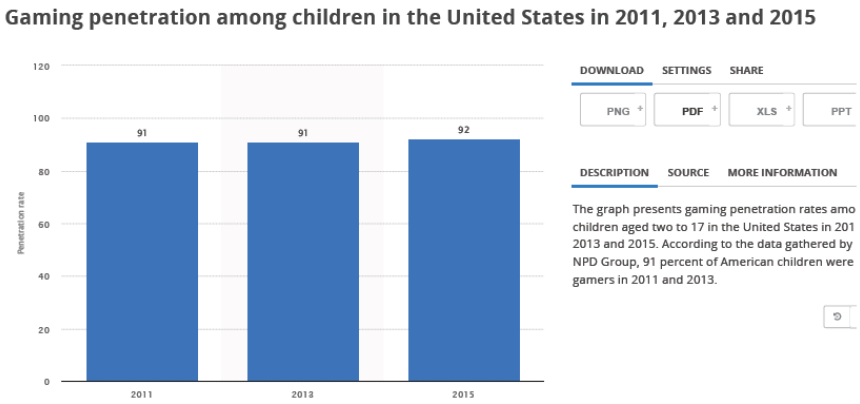
Researchers have studied the effectiveness of serious games in providing basic and refresher training. French Military Medical Services sponsored the development of a series game, 3D-SC11, for Tactical Combat Casualty Care (TCCC) training programs, Sauvetage au Combat de niveau 1 (SC1). Planchon et al. (2018) find that additional training with 3D-SC11 has improved the performance of the soldiers. The authors conclude that 3D-SC11 could be an efficient tool to train soldiers in life-saving interventions, and may also be used as a cost-effective and engaging way to training the wider civilian public. Li et al. (2017) investigate the effectiveness of using a serious game to train new workers on complex manufacturing tasks as compared to traditional training, and find that training game is more effective for learning procedural knowledge than the traditional paper manual based training, but there is no significant difference in terms of factual knowledge. The primary research question of Nicolaido et al. (2015) is "What are professionals' perceptions of the potential of the Virtual Emergency Telemedicine game for training people involved in the assessment and management of emergency cases?". The feedback from the professionals are mostly positive, supporting the use of the game as an effective training tool, despite identifying some confusing features.

Serious games have also been used as an assessment tool. Shimic et al. (2015) present the development and application of Serious Gaming Interactive Questions (SGIQ) to assess the readiness of doctors, surgeons and other medical staff for AMET's (Air Medical Evacuation and Transport) complex task of transfer of injured or sick patients from the place of accident to the hospital or between hospitals. The SGIQ-based assessment is designed for evaluating three kinds of skills and knowledge necessary for successfully participating in AMET missions: (1) Knowing and fast finding of equipment and medicament in the aircraft; (2) Monitoring the patient state during the transport; and (3) Knowing the procedures and actions appropriate to the patient's state

and conditions. The difference between SGIQ and training games is that the targets of SGIQ are high-skilled adults who already have sufficient domain knowledge and a lot of experience.

3.3 Gamification

Gamification is the process of taking something that already exists – a website, an enterprise application, an online community – and integrating game mechanics into it to motivate participation, engagement, and loyalty. Gamification takes the data-driven techniques that game designers use to engage players, and applies them to non-game experiences to motivate players. Figure 1 shows that 92% of the children between 2 to 17 in the United States played video games in 2015. As young adults, they would most likely be more actively engaged in activities with gaming elements.



Source: https://www.statista.com/statistics/274600/gaming-penetration-among-children-in-the-us/

Figure 1: Gaming Penetration among Children in the United States in 2011, 2013 and 2015

Kerfoot and Kissane (2014) and Mokadam et al. (2015) find that the introduction of gamification, focusing on technical and cognitive skills, have encouraged the use of simulators by surgical residents, capitalizing on the competitive nature of the trainees. Both studies conclude that gamification helps to achieve demonstrable improvement in the performance of the residents.

Gamification is not only effective for those with high "internal" motivation, such as surgical residents, it is also an effective tool for other vocational trainings. Xiang et al. (2014) conduct a series of beta tests of applying gamification to help students to engage in learning of Info-Communication Technology skills. The study finds that the use of games help improve students' learning up to 80% with a mean of improvement of 34.4%. Hoglund (2014) provides a good literature review on roles of games and gamification in employee engagement, motivation

and interest for training activities. The paper also discusses the challenges for applying games and gamification in professional and vocational training. Foss et al. (2007) discuss the development of an e-learning programs based on dynamic simulators and features of computer games to meet the training needs of automation technicians, including practical experience and a good understanding of basic principles of automation and control systems in order to be able to maintain, trouble-shoot and solve problems related to automation and control of industrial processes.

Traditional astronaut training provides "the required knowledge, skills and attitude to execute the complete set of tasks and experiments to be performed in a short mission" (Cornelissen et al., 2012). However, with longer missions, such as long term stay in ISS, trips to Moon and Mars, there might be a need for refresher of pre-mission training and new training materials during the mission. The results of Cornelissen et al. (2012) indicate that introducing gaming aspects in the training of astronauts for long duration missions help to overcome the possible decline in motivation and loss of interests in self- studies.

4. Advanced Technical Programs in Healthcare Training

Training of healthcare workers is critical for not only saving human lives, but also in maintaining the quality of lives. Two critical aspects of healthcare are surgery and nursing. In this section, we describe the advances made in training of surgeons and nurses across the world.

4.1 Surgeon Training

Background

Surgeons have patients' lives in their hands. They are trained to be precise with their hands, be able to make decisions in stressful situations, and act as the leader of the multi-disciplinary team in the operating theater. Their performance depends on their skills, their acute situation awareness, and their ability to work and lead others on their surgical team. The outcome of their performance have direct and immediate impact on the welfare of their patients. When a mistake is made and the patient dies on the operating table, the outcome is not reversible, and there is no chance for correction. Although air traffic controllers and surgeons do not share many common technical skills, they both operate under pressure, and need to make quick decisions that may lead to dire

consequences. Naturally, the training programs for air traffic controllers and for surgeons do share many similarities in terms of how technical skill trainings are be provided. This section reviews some current practices in surgical training programs that may have implications for ATC training.

Surgery involves many tasks, which require co-ordination of different skills and performance under high pressure of either time bound routine, emergency setup or immediate complications. The environment is usually stressful with multiple events occurring at the same time and multiple individuals carrying out interdisciplinary tasks in a coordinated fashion. Usually the surgeon is the leader and needs to perform the surgery, and also co-ordinates with others and deals with any issues that arise during the surgery at the same time. Consequently, the skills required for avoiding errors that are reliant on skills, rules or knowledge in these procedures belong to a broad spectrum. These skills are psychomotor, cognitive and interpersonal in nature. As mentioned before, surgery takes place in stressful environment and the surgeon needs to master not only technical skills, but also combine these with multitasking and potentially keep some attention span for room for error, which may be costly.

Surgeons are trained in stages and with different training methods: classroom based knowledge, observation of video demonstrations, and ex vivo and in vivo animal models, on job hands-on training, traditional internship/residency, and increasingly simulations and virtual reality based training. Each area of surgery may use the same or different order from that mentioned above. The protocol is also different by country. However, what is common to training of surgeons across all geographical locations and fields is that there are only a limited number of hours that can be spent on training with traditional methods. Consequently, there is an increased focus on immersive virtual reality in surgeon training. Immersive virtual reality allows surgical trainees to experience various “realistic” scenarios in operating rooms, and to develop surgical skills, expertise and subconscious awareness of conditions in the OR (operating room) under pressure, which enables a surgeon to achieve automatic cognitive processing of various roles needed to be performed when the real situation so demands the same

Understanding human behavior is critical for establishing and developing an effective training program. Schneider and Shriffin (1977) and Shriffin and Schneider (1977) find that there is automatic search for familiarities which leads to quicker response in subjects. Logan (1990)

finds that exposure to and quick reflection from experiences in the past lead to quick retrieval of steps needed to act when situation demands the same. These findings suggest that automaticity may be speeded up by practice and that repetition leads to faster retrieval of skills and knowledge. In the case of surgeons, surgical skills can be improved by "practice and practice". Immersive simulation, such as in the form of virtual reality, therefore, provides a great platform to allow surgical trainees to practice certain skills repetitively to achieve automaticity, in layman's words, "they can do it in their sleep".

Rasmussen (1983) classified human behavior as based on knowledge, rules and skills.

- Skill-based behavior is related to sensory-motor related and are mostly integrated within the human psyche due to repetition. . Hence, surgery simulations, box trainers and virtual reality (VR) simulators are effective tools to address basic surgical skills (Dargar et al., 2015).
- A collection of skill-based behaviors is stored in the mind of the person in the form of a rule. When a particular scenario appears, the rule would dictate certain actions by the surgeon. This is referred to as the so-called rule-based behavior.
- During unfamiliar circumstances such as in emergency or crisis, a surgeon must utilize the concept behind the rules and skills learned from previous encounters and intuitively extrapolate these to derive optimal outcome. This involves tertiary level of thinking and falls under knowledge based behavior. VR simulators that combine skills and processes involving different rules under unfamiliar circumstances for trainees, and introduce complications can be useful for stimulating and assessing a trainee's knowledge based behavior. The overall objective of simulation based training in surgery is to enable a surgeon to reach the level such that they can intuitively make the right decision and use the correct skills under unfamiliar circumstances based on knowledge accumulated from various prior scenarios (Gallagher et al., 2005).

Figure 2 illustrates Rasmussen's behavior model graphically.

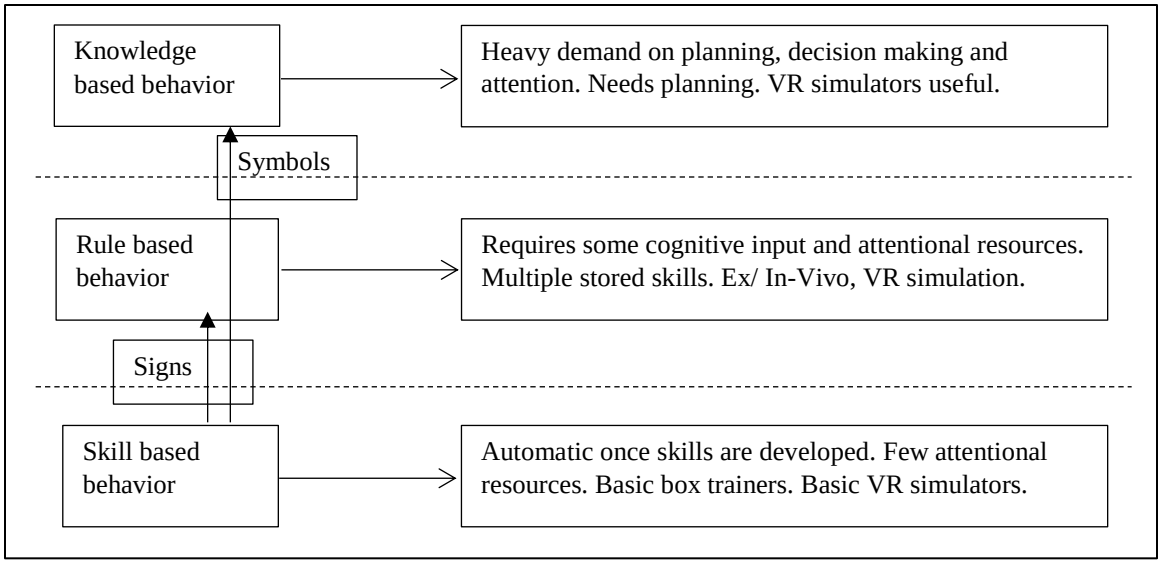


Figure 2: Broad outline on key learning about translating Rasmussen's (1983) model of human behavior into development of training methods

Wentink et al. (2003) proposed a framework for creating laparoscopic surgery protocol based on Rasmussen's model, considering all three levels of human behavior. They observe that most of existent VR simulators concentrate on skill-based behavior only, thus further development in training programs and associated technologies should focus on rule and knowledge based behaviors.

An effective training program that utilizes Rasmussen's model should endeavor to first develop trainees' familiarity with their working environments through skill training, then advance their understanding of different tools and techniques for various objectives based on the concept of rule based learning, and create complex scenarios to accomplish knowledge based learning (Drivalou and Marmaras, 2009). Elaborate planning to develop training protocol at all three levels is essential since potential mishandling of tools and techniques (skill training) may lead to complications in surgery, missing out on a step in the surgery procedure may both botch the entire operation, and hesitation or anxiety in the presence of unexpected event such as more than expected lesions in surgery may lead to catastrophic consequences (Dankelman et al., 2004).

4.2 Development of training program

Gallagher et al. (2005) stress on the need to plan the education and training program before introducing VR simulators. They also stress on the presence of tractable performance metrics relevant to the surgical task being performed. The study stresses on staggering the training program rather than creating a single program to imbibe all the content. The Key finding from the paper is that training program needs to be carried out in small achievable modules for better assimilation by trainees.

Molduvanu et al. (2011) tend to agree with the point that if surgeons are exposed to some sort of warm-up training, then they are more likely to perform better in surgery. This observation was based on observation by experts on two sets of individuals – those who were exposed to warm-up exercises and those who were not. However, this research did not find the difference to be statistically significant, possibly because of a small sample size (20 patients). Yiannakopoulou et al. (2015) agree with the view that the qualitative evidence of superior performance due to VR simulators needs to be backed by quantitative data, which is currently scarce in surgery. Key finding from the paper is that while there is a lack of data until early 2010s, evidence from then does not support the hypothesis that VR based simulations improve surgery skills. However, as more and more studies are carried out especially after 2015, there is increasing evidence in diverse fields of surgery that VR based simulations make trainees equal to or better off than their counterparts who do not receive such training.

Villard et al. (2014) develop a framework of working on a physical mannequin coupled with haptic feedback and pseudo-ultrasound generation to simulate a real time computerized immersive virtual environment for training in interventional radiology. Metrics used for assessment measure skill level performance measures such as areas hit, that were required, areas that were unnecessarily hit and overall time taken. Key learning from the paper is that with increased regulatory restrictions on working hours, patients and problems with animals, fully immersive VR simulators are rapidly becoming relevant to training.

In line with the above Palter et al. (2011) find using an assessment rating scale that deliberate practice group of surgical residents performed superior in OR as compared to the control group, thus underscoring the importance of VR curricula in laparoscopic surgery. This surgery

involves making small cuts to perform precision jobs in human body. The findings of Huber et al. (2017) seem to resonate with this observation. These authors combine head mounted displays (HMDs) with VR laparoscopic simulator (Figure 3) and asked 10 members of the surgical department to first use the simulator and then report their "exhilaration scores". The study finds that participants spent more time on the simulator with HMD and also felt more exhilaration. Key learning from this study is that VR curricula produce desired results of superior learning and performance. Further, greater the immersion of the VR simulator, greater is the willingness of participants to learn and more is the time spent in learning. Also participants feel greater sense of accomplishment with the VR simulator with HMDs.

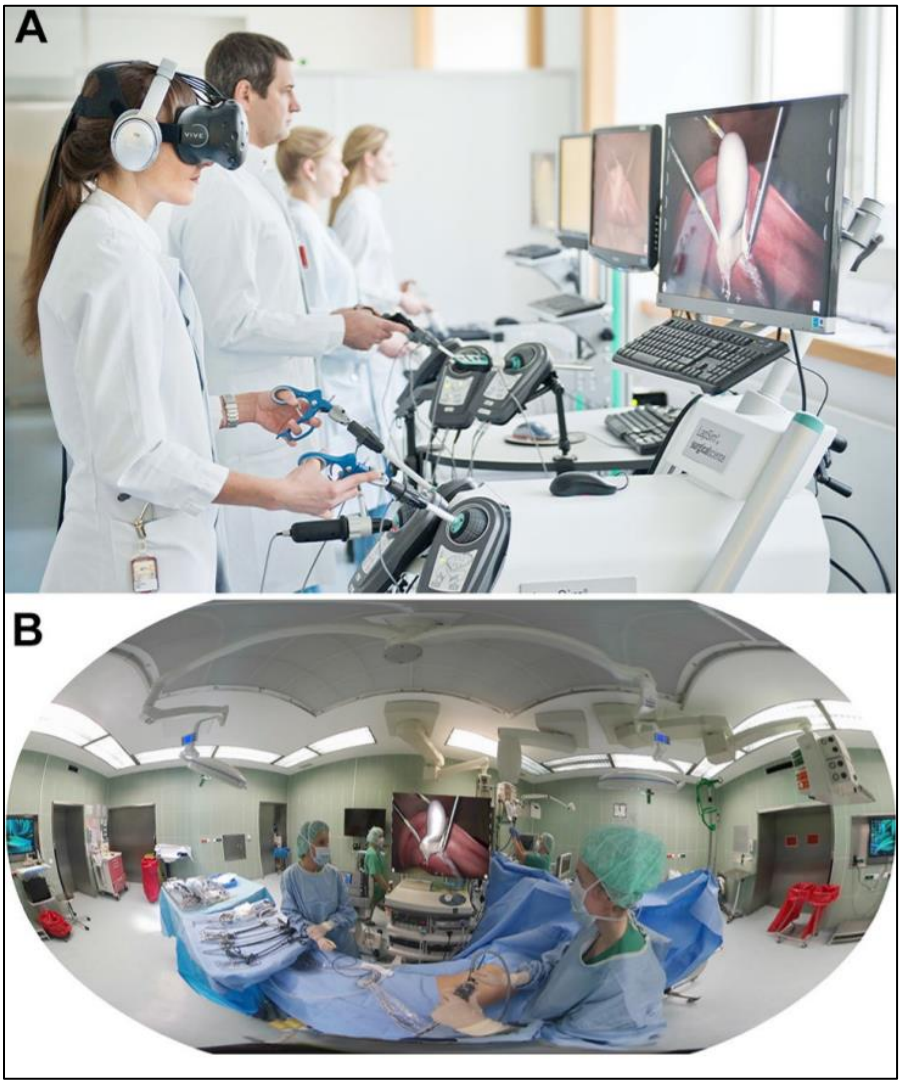


Figure 3: Use of head mount displays (HMDs) with virtual reality (VR) simulation for Laparoscopic surgery (courtesy Huber et al., 2017)

Top A: external view of the setup with regular and IVR simulations.

Bottom B: personal 360° immersive operating room view of the study participant during IVR session

Crochet et al. (2017), Gallagher et al. (2018) and Torkington et al. (2001) report the use of minimally invasive surgical trainer (MIST) VR for laparoscopic surgery. All of these studies report measuring the performance of surgeons with respect to assessment scores based on precision of incision, number of movement, path of movement, path length and accuracy, idle time, total time, time spent on incorrect path and respect for tissue.

El-Beheiry et al. (2017) find that introducing competition into training increased the voluntary usage of the laparoscopic simulator. The study had two groups – control group without competition and the competition group. The latter group overwhelmingly did not want competition to be a part of the simulation-based training in the beginning. However, as the leaderboard was posted biweekly and monthly for the fastest peg-transfer times achieved by residents, they seemed to be motivated. Consequently, the competition group spent significantly more time at the simulator and also achieved better results at peg-transfer. Key learning from this paper is that introducing competition motivates trainees to learn better.

Roberts et al. (2017) use a passive haptic VR simulator called (ArthoS) in arthroscopic procedures for either knee or shoulder (Figure 4). Performance of 30 participants was assessed using a validated global rating scale on the blind videos of the performance. After the procedure, the participants were surveyed as well. The paper finds the overall VR simulator is effective in meeting its performance objective. However, the participants felt that haptic feedback would offer further simultaneous improvement. Key learning from this paper is that VR simulators in surgery are becoming more mainstream and accepted readily. These are an important addition to conventional methods of training in surgery.

Figure 4: Artho MIST VR used in Virtual Reality simulation for shoulder and knee

Perry et al. (2017) surveys and identifies curriculum used in training and learning in dental education. Traditional curricula include imparting a preclinical basic science foundation with clinical experience pushed back for the later years. The focus is on lectures that are divided into different subjects. On the other hand, integrated curricula include traditional lectures blended with problem based and case based learning. Addition of many complex, time consuming dental procedures have made these curricula more demanding leaving little or no time for students to reflect and analyze. It is for this purpose, a number of curricula have explored the use of VR simulations and found that their use affects the performance positively (Imber et al., 2003). Use of VR simulations also help identify students how are less likely to perform well in surgery and helps instructors create additional resources to provide support and feedback to such students (Gottlieb et al., 2011). Gottlieb et al., 2017 additionally investigate simulation parameters such as exam scores, self-evaluations and time to complete evaluation served to predict dental students' preclinical performance. They found that advanced simulation scores may help identify students' weaknesses early and help them cope with these in a constructive way. Key finding from this paper is that training in dental education takes place in traditional curriculum and also uses blended VR simulated environment to provide hands-on learning. Assessment is through tests that are theoretical as well as simulator surgeries.

Gallagher et al. (2005) note that VR training is successful if the curriculum is well thought out with a focus on needs of the participants (such as skill development/ improvement). Training is likely more effective if the schedule is based on regular intervals rather than amassed into a focused session (Figure 5). Further, high fidelity simulations provide greatest skill transfer to

participants but are also considerably more expensive. Their proposed template for a successful curriculum consists of:

- Teaching of relevant knowledge
- Step-by-step instruction of tasks to be performed in the procedure
- Defining and illustrating common errors
- Testing of the knowledge imparted
- Training technical skills on the simulator
- Immediate feedback when error occurs
- Summative end-of-the procedure feedback
- Repeated trials as necessary

Some training strategies discussed by the paper are:

- Shaping:
    - Successive approximations of the desired response pattern are reinforced until the final outcome is achieved.
    - Start with simple hand-eye co-ordination tasks
    - Increase the level of complexity on the tasks performed for gradual skill development
    - Transition from one level to the next should be smooth
- Fading
    - Involves giving trainees a number of clues and strategies in the beginning
    - As tasks become more complex, very few to no aids are provided to the trainee
- Backward Chaining
    - For complex tasks with high failure to success ratios, this strategies is commonly used in surgery
    - It breaks down tasks in multiple parts
    - It addresses the last part first and help build successively each of the previous step
    - This enables the trainee to visualize end at first, which in turn reduces trainee frustration.
    - As each previous step is added, trainees get more and more proficient at the end

One word of caution provided by authors is that simulators can teach and reinforce bad behaviors. This can be dispensed by making sure experts go through the simulation before exposing trainees to training curriculum. To summarize, this paper provides contextual framework for training and simulation based learning. It also discusses different strategies for training curriculum.

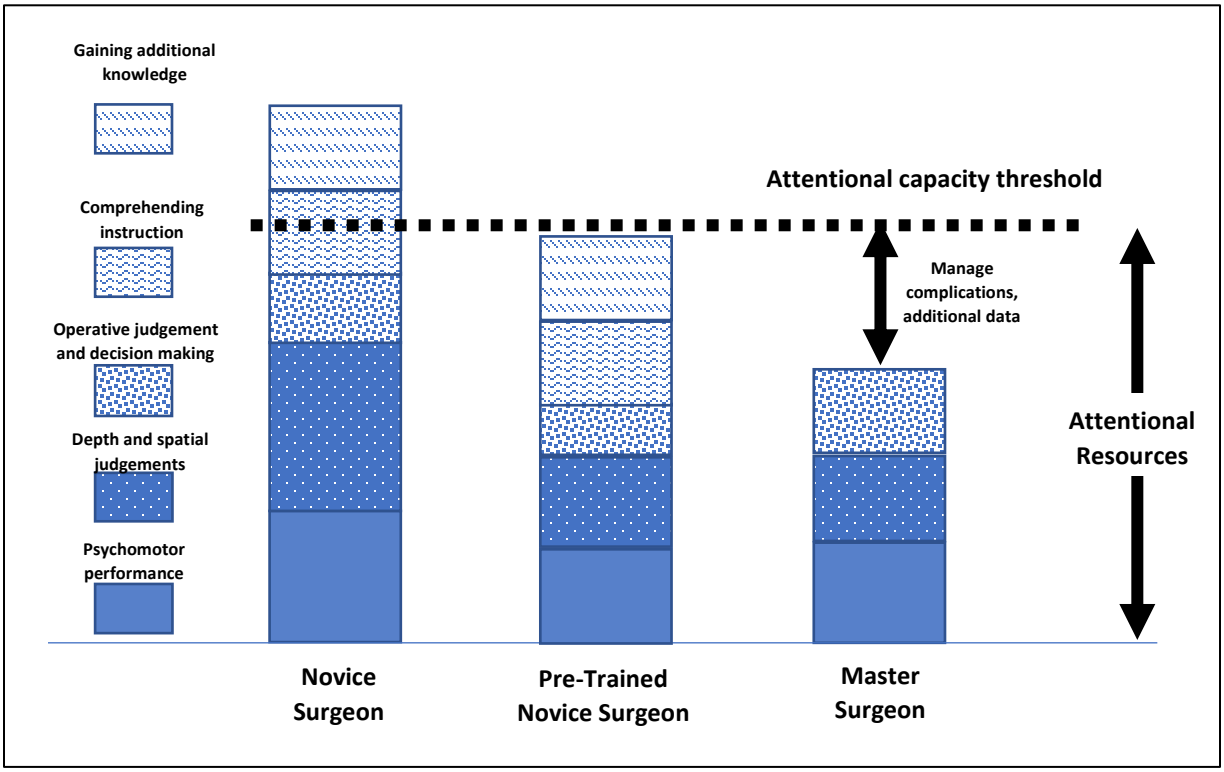


Figure 5: Hypothetical attentional resource benefits of simulation training adapted from Gallagher et al. (2005)

Seabrook et al. (2005) carry out three different experiments across different age groups and different time lags to investigate whether it is better to make mass presentations in a teaching curriculum or it is better to have lessons that are distributed with different lags. The study concludes that irrespective of the age groups of the learner (s), retention of concepts is greater if the presentations contain a lag. However longer lag effects do not produce significant difference in learning. Key learning from this paper is that spacing presentations improves retention of content in the memory of adults as well as children of all ages. Furthermore, retention improves with the duration of the gap. However, very long spacing times are not very beneficial, implying that the benefits of spacing time duration level off with increasing time.

Ploch et al. (2016) review the use of 3-D printing in surgical planning, practicing and teaching to create devices are resemble human organs for applications in cardiovascular, neuro, orthopedic and plastic surgery. These models are deformable and resemble the human organs more closely than the hard-plastic models which are conventionally used. The authors demonstrate how to fabricate 3-D printed models that are easy to produce, durable, softer and cost effective. Key Learning from this study is that 3-D printing with right materials can produce objects relatively cheaply and which resemble the actual parts to help trainees practice and gain confidence with their required job skills.

Vakharia et al. (2015) review the work of 31 different studies that use 3-D printing in the field of neurosurgery. Many of these studies actually use these models with simulations for training in surgery. These models help trainees and surgeons practice parts of operation pre-operatively. The main finding of the paper is that by replicating the patient specific pathologies, 3-D models provide an extremely accurate perspective with which surgeons can overcome the challenges of shorter working hours, growing number of techniques and increased sub-specialization.

Dawe et al. (2014) review transferable skills to patient based setting from simulation based surgical training. The paper explores as many as 19 different procedures and included both interns and trainees in surgery. Following were the areas studied with main findings:

- Overall findings

  Many studies have reported an overall assessment by either patients or assessors to evaluate the overall performance using some kind of a score (a global rating scale based either on mean or median score). The findings were based on a number of laparoscopic studies involving:

  - LapSim VR simulator, (surgical residents in Sweden), (Ahlberg et al., 2007)
  - Limbs and Things Laparoscopic simulator (interns in USA), (Banks et al., 2007)
  - FLS training Box simulator and Lap Mentor VR simulator (general surgery and gynecology residents in France), (Beyer et al., 2011)
  - Lapsim VR simulator (surgical trainees in Australia), (Cosman et al., 2007)
  - FLS Training Box simulator (gynecology residents in USA), (Gala et al., 2013)

- Lapsim VR simulator (surgical residents in USA), (Hogle et al., 2009)
- Lapsim VR simulator (gynecology and obstetrics in Denmark), (Larsen et al., 2009)
- FLS Training Box and Lapsim VR simulator (general surgery residents in Canada), (Orzech et al., 2012)
- FLS training Box simulator (general surgical residents in Canada), (Sroka et al., 2010)
- MIST-VR simulator (senior surgical residents in USA), (Van Sickle et al., 2008)
- TEP Hernia task trainer (general surgery residents in USA), (Zendejas et al., 2011)
- ProMIS surgical hybrid simulator (medical students in Switzerland), (Franzeck et al., 2012)
- LapSim VR simulator (general surgery residents in Canada), (Palter and Grantcharov, 2012)
- LapSim VR simulator and FLS training box (general surgery residents in Canada), (Palter et al., 2013)

Key findings and learnings:

All these studies involved a wide variety of procedures and tasks and mainly found that there was improvement in scores, improvement in time of the procedure, better transfer effectiveness ratio, better dexterity, fewer suturing errors and fewer complications. Only a few of these studies found no statistical difference between simulation based training and conventional curriculum based training.

- Endoscopic procedures

  A number of studies involving endoscopic procedures have reported a higher overall performance accuracy using colonoscopy procedures for participants who have received simulation training.

  - Low fidelity transnasal fiberoptic flexible laryngosocopy simulator (medical students and junior residents in Canada), (Deutschmann et al., 2013)
  - GI mentor VR endoscopy simulator (residents in internal medicine in Austria), (Ferlitsch et al., 2010)

- Endoscopic sinus surgery simulator (otolaryngology residents in USA), (Fried et al., 2010)
- PelvicVision VR simulator (urology residents in Sweden), (Kallstrom et al., 2010)
- Nasal model endoscopic simulator (medical students in USA) (Ossowski et al., 2008)
- AccuTouch VR colonoscopy simulator (general surgery and internal medicine students in Canada), (Park et al., 2007)
- URO Mentor VR endourological simulator (interns in The Netherlands), (Schout et al., 2010)
- GI Mentor II VR endoscopy simulator (first year gastroenterology fellows in USA), (Sedlack, 2007) and (medical residents in Japan), (Shirai et al., 2008)
- Colonoscopy simulator II (6 fellows and 5 residents in South Korea), (Yi et al., 2008)
- Clinical plus simulator training for IG1, IG2 and CG (medicine and surgery residents in Germany), (Ende et al., 2012)
- Endoscopic sinus surgery simulator (otorhinolaryngology residents in USA), (Fried et al., 2012)
- Olympus colonoscopy simulator (novice colonoscopists in UK), (Haycock et al., 2010)

- Other Procedures
  - VIST VR simulator (cardiology trainees in Canada), (Bagai et al., 2012)
  - Eyesi VR surgical simulator (third year ophthalmology residents in USA), (Belyea et al., 2011)
  - Sawbones arthroscopic knee simulator (junior orthopaedic surgeons, diagnostic knee arthroscopy in UK), (Howells et al., 2008)
  - VIST VR simulator (first year general surgery registrars, superficial femoral artery angioplasty in Ireland), (Hseino et al., 2012)
  - Eyesi VR surgical simulator (ophthalmology residents in USA), (McCannel et al., 2013)
  - Synthetic abdominal wall model (gynaecology residents in Canada), (Palter et al., 2011)

- Eyesi VR surgical simulator (opthalmology residents in USA), (Pokroy et al., 2013)

Almost all these studies reported better performance time, as taken in minutes or seconds to conduct a patient based assessment procedure or task. One of these studies (Hseino et al., 2012) reported better scores in terms of "ability to complete" the case during the procedure. Studies by Ahlberg et al., 2007; Cosman et al., 2007; Ferlitsch et al., 2010 and Van sickle et al., 2008 reported fewer errors by participants after successfully completing simulation based training.

Overall conclusion drawn from these studies indicates a benefit for both trainee physicians/ surgeons as well as patients with respect to safety, efficiency and cost savings. All studies have pointed to either greater or at least equal performance as without simulation-based training.

Conclusion and Key learnings from the training of surgeons:

- Situation awareness is a key indicator of measure of person's level of responsiveness and preparedness for a complex situation.
- Role play and responses to hypothetical situations are useful in developing skill based behavior toward desired situations.
- Elaborate planning and insight into situation is a key to develop knowledge based and rule-based curricula for trainees.
- Training programs that are created in chunks are better assimilated.
- VR simulators are becoming very popular for surgical training.
- VR curricula produce desired results of superior learning and performance.
- Introducing competition motivates subjects to learn better.
- Assessment is through tests that are theoretical as well as simulator surgeries.
- There has been much progress in the development of virtual reality based devices for enhancement of surgical skills. These devices target skill based, rule based and knowledge based behavior of users.
- Immersive virtual reality is essential to gain mastery.
- Immersion can be measured by presence surveys such as developed by Witmer and Singer (1998) and used by Silva et al. (2016).

- The path towards complete immersion is tough and the main challenge is cost. Higher immersion needs greater fidelity and comes at the cost of more complex systems in terms of haptic feedback as well as computational power as described by Caird (1996).
- The results are encouraging with a surge of new technology flooding the markets in all areas of determining presence –haptic devices, motion/ control devices and trackers and HMDs.
- It is already possible to generate interactive simulations with deformable organ models which can be easily printed using 3-dimensional technology. More complex models that provide feedback can also be generated.

4.3 Nursing

Simulation has been extensively used in training nurses, and has proven to be more effective than the conventional form of education (Alinier et al. 2006). High fidelity [2]training involves training near realistic scenario (computerized human patient simulators). Intermediate fidelity training is defined as the training under semi-actual situation in which the trainees are exposed to some of the elements of the real-world scenario (such as mannequins, video games, medical devices and some exposure to invasive and non-invasive techniques). Low fidelity training is defined as training with task trainers with a traditional protocol review and no simulation (Harder, 2008). Cant and Cooper (2017) report 25 different reviews of simulation-based training from 2010 to 2015 for pre-licensure nursing students reporting that although no quantitative overall effect was derived, the self-efficacy reporting by the students suggested an increase in overall confidence, skills and critical thinking ability. The review does note some gaps in evidence of effects that need to be addressed.

The sample size in such studies for practical reasons can be small to large (from 10 to 1000) depending on the number of participants and their willingness to participate in such a study. In the Alinier study, students were randomly chosen in control group and experimental group. Both the groups were administered the traditional training using clinical methods, administered a pre-test, and a post-test and surveyed using a questionnaire. The training was over a period of two days. However, the study was carried over a period of 2-years. An assessment of trainees was carried

[2] Appendix A defines and discusses the term “fidelity” in training.

out after 6 months. The participation in the experimental group was voluntary and members were exposed to intermediate fidelity training in addition to the traditional style. This training involved participation in teams of 2. The study finds an increase in the performance of individuals that were exposed to simulation-based training in an objective structured clinical examination (OSCE). This increase was statistically significant. An OSCE is composed of 15-20 stations and students spend 3-10 minutes at these stations either practicing a skill they learned or writing a response to a question.

Harder, 2008 reviews 61 different research papers in the field of medical education (nursing, medicine and interdisciplinary) that report the use of high fidelity simulations. Nearly all these studies employed some form of assessment (either quantitative such as OSCE or qualitative such as surveys and self-reported change in confidence). Some other training methods reported in the literature were as follows:

- standardized patients
- traditional psychomotor skills
- laboratory sessions with task trainers,
- computer-based programs, and
- lecture classes

Only three studies reported no difference in the learning of students with simulation-based training. The conclusion of this study is the growing use and popularity of human patient simulators in both nursing and other medical fields in high fidelity simulation set-up.

As the popularity of high fidelity simulation has grown, increasing research is on observation of the simulation and debriefing after the simulation. Johnston et al. (2017) find that debriefing of the group that participated in the simulation, followed by debriefing based on the transfer of learning principles may be a useful way to learn from the simulated experience. In another study by Booth et al. (2017), the reflections of participants in the simulated study reveal that participants “greatly” grow in confidence and skill levels to connect with and help mental health patients. This study employed 12 baccalaureate nursing students involved in six 10-hour shifts in a clinical set-up and six 2-hour simulations. This is a qualitative study which bases its findings on document study and focus group interviews.

Another study in nursing by Bliss and Aitken (2018) finds that the use of simulation enhanced their skills in recognizing and responding to patient deterioration in order to escalate the issue for prompt care of patient in need of critical care. This study also employed the use of qualitative data following semi-structured interviews. Results were collected using anecdotal evidence based on experience of the registered nurses. Students' attitudes and perceptions also play a critical role in learning from simulation-based training. In another study by Gharaibeh et al. (2017), previous experience in intensive care helped the nursing students have a better attitude toward simulation. This study uses multiple regression technique on a sample of 413 nursing students to find the prior perceptions and experiences were the only significant factors in influencing their attitude toward simulation-based training.

The study by Biglino et al. (2016) uses intermediate fidelity simulation based on 3-dimensional models of congenital heart disease during a training course of nurses and provides recommendations to optimize their use for the purpose of training. The main findings of this study were in line with the existing research that the simulation was well liked by the nurses. The subjects were shown many different three-dimensional models following a lecture on the process of making the models. Participants were asked if the models were more illustrative than the actual drawings used in the lecture. Written feedback and survey data were used to conclude that the models were superior to the traditional lecture set-up.

Another review study in surgery by Langridge et al. (2018) demonstrates that the use of three dimensional models obtained due to 3-D printing in medicine especially in neurosurgery and otorhinolaryngology. The usage of 3D printed models in training and education has also been used in anatomical education and preoperative planning. This shows better results when compared with educational methods used in a traditional environment. The use of 3D models greatly enhanced surgeon skills and also improved the outcome of the surgery for the patients.

Key learnings from Nursing:

- Laboratory task trainers are increasing in popularity and complexity in the area of training nurses.
- Practice of skills with mannequins that provide auditory/ haptic feedback is gaining ground.

- Virtual reality improves confidence, skills and the ability to solve problems under pressure.

5. Pilot Training

The Federal Aviation Administration (FAA) acknowledges that pilot currency and refresher training are critical to the safety of flight in both visual and instrument conditions (LeRoy Thomas, 2018). In addition to mastering the psychomotor aspects of piloting an aircraft, refresher training includes the assimilation of new information as well as the reinforcement of previously learned knowledge.

The following provides a summary and literature review of some of the methodologies and lessons learned from pilot training industry from academic and trade journals. These articles can be divided into:

5.1 Innovative Learning Methods

There are many studies that recommend departure from old classroom materials for pilot training to active and innovative methods. The following presents some of these studies who used statistical analyses to identify more efficient and improved training methodologies:

Caro (1973) discusses the importance of flight simulators as primary pilot training vehicles. The study emphasizes on metrics that have been instrumental in pilot training such as minimizing over-training, effective utilization of personnel, use of incentive awards, peer training, and objective performance measurement.

Hamman (2004), discusses the relevance of aviation and health care industries. The paper concludes that teams are responsible for safety and accordingly team training should receive equal importance as technical materials. Therefore, it is recommended that team training skills be included in curriculum and team training skills must share equal importance with the technical skill requirements.

Hales and Pronovost (2006), review practices in aviation, health care and manufacturing industries where the welfare of a human being is at risk, and recommend checklists to reduce human errors. They conclude that using simple tools such as checklists increases safety and better work environment.

Smith, et. al. (2010), analyze more than 2000 airline pilots and statistically identify best performing pilots as follows:

- Had flight instructor certificates
- Graduated from collegiate flight programs
- had between 500 and 1,000 pre-employment flight hours

McClernon (2011), recommends including stress exposure training as part of their overall simulator curriculum. They statistically concluded that pilots with this training flew much smoother when faced with a situation than those who were not exposed to this training.

IFALPA (2012) Pilot Standards held a workshop in 2010 to help pilot training schools with guidelines recommending best practices. Some of the recommendations from this workshop include:

- Flexible curriculum – the curriculum should avoid one size fits all and provide flexibility to students and instructors to adapt the content.
- Training vs. Checking – During the recurrent training, less checking and more training is desired. The emphasis must shift from checking pilot skills to training pilot skills.
- Scenario based training – The emphasis here is on critical thinking. The goal here is to accelerate higher level of decision-making skills.

Marvin and Roth (2014), recommend the notion of team training. They conclude that pilots who critically assessed the practice of other pilots in pre-recorded videos felt empowered by transferring classroom instruction to the workplace. They conducted this study to help an airline improve their pilots training system. They recommended recording pilots engaged in flying using simulators and make these available to all trainees to re-create real workplace examples. The trainee pilots are asked to identify ways to improve as an individual and as a team.

Sommer (2014), identifies that current airline pilot training consists of basic training and type-rating. They include training methods comprised of classroom instruction, computer-based training and practical training, in either the aircraft or a simulator. Pilot training not only includes technical and procedural instruction, but also training in non-technical skills like crisis management, decision-making, leadership and communication. The study then draws some parallels between surgeons and pilots. They are both under high physiological and psychological stress, operating expensive equipment, and the ultimate cost for error is measured in human lives.

The study then recommends lessons learned from surgery training to be equally applicable to pilot trainings including:

- Intensive selection of candidates before training;
- Technical, procedural and interpersonal training;
- Detailed program/syllabus;
- Objective evaluation of performance;
- Simulation devices;
- Life-long recurrent training and qualification

Hong, et. al., (2016), conduct a survey of 120 Asian student pilots to assess their training and perception of their institutions. Their statistical analyses reveal that student pilots who lack confidence with respect to their knowledge during flights; fail to recognize the importance of maintaining relationships among supporting staff such as air traffic controllers, mechanics, and others involved in the flight process. They recommend the training centers to foster awareness of interaction between human factors and other aspects of aviation safety.

LeRoy Thomas (2018), attempts to identify the impact of different training methods on pilot learning. They use three methods of training on 62 non-current instrument rated pilots as follows:

- Passive learning – just reading the materials. Conceptual knowledge is presented to the learner in the form of facts and theoretical training.
- Active learning involving some degree of simulated flying approaches. This type of learning will engage the learner more directly in the learning process

- Experiential learning involving scenarios and requiring decision makings.

Before and after the trainings, each pilot took knowledge tests and self-efficacy questionnaires as a measure of training effectiveness. The study then utilized these measurements to conduct statistical analyses on which methodology worked the best. They statistically concluded that:

- Increase in knowledge scores between pre-training and post-training independent of training types.
- The increase in knowledge was significant between the passive method of training (just reading) and experiential learning which includes full scenarios and decision making.
- Using active learning through personal computer-based flight scenarios proved to be the most effective method of refresher training for instrument rated pilots who are not instrument current.

5.2 Pilot selection candidates

Several studies highlight the importance of selecting pilot candidates. They conclude that these selection processes should consider both physical and mental capabilities of the candidates. Some of these studies include:

Burke et. al. (1997) and Pulos (2009), use a pool of air force pilots across several countries to identify the selection process for pilots. They identified that pilots' psychomotor and high cognitive ability are dominant factors in successfully completing their trainings and performance in later stages of professional career. They conclude that neuroticism and its facet anxiety would be negatively related to training success, and that extroversion would share a positive relationship with training success in military aviation.

Smith, et. al. (2010), analyze more than 2000 airline pilots and statistically identify best performing pilots as follows:

- Had flight instructor certificates
- Graduated from collegiate flight programs
- Had between 500 and 1,000 pre-employment flight hours

European Pilot Training ECA(2013), reviews the pilot training in Europe and the factors that make a good lifelong commercial pilot. It starts with a selection of good candidates who want to become pilots. A major pre-requisite is candidates with ability to think outside the box in order to deal with unexpected and undefined events. Some of the required metrics include, being healthy, self-motivated, have reasonable psychomotor skills and emotional stability. The current good performing pilots must be part of selecting committee. The training programs and curriculum are relevant, innovative and data driven.

Turner (2014), outlines methods for selecting pilots whose first language is not English. The study proposes and applies a methodology to select pilots for Chinese airlines. It concludes that culture plays an important role in pilot training. The two notions of “interdependence” and “independence” are contrasted between the two ways of “being a person” in European-American versus Asian cultures (including the Chinese culture) in pilot training.

Roth (2015), conducted a study in in Australia and Canada on how flight examiners reason flight and which methods they use when assessing the competencies of pilots for accreditation purposes and type-rating training. Their study includes 23 flight examiners from 5 different airlines. These examiners typically use documentary methods of interpreting their observations. They identified that the metrics used are different between senior and junior flight examiners. They concluded that senior examiners (more than 2 years of experience) use more subjectivity and are more forgiving for small errors in their assessments while less senior flight examiners (with less than 2 years of experience) focus more on individual facts and discuss small errors without looking at the overall competency of the pilot.

5.3 Recurring training

A large number of studies have recommended ongoing and recurring training for pilots. These studies show that experienced pilots become overconfident and take unnecessary risks. Some of these studies include:

Bazargan and Guzhva (2012), look at the combination of age, gender and experience of GA pilots. They conclude that while female pilots with the same level of experience as male pilots are marginally safer, this difference starts to diminish over time. They conclude that pilots with

low number of flight hours (less than 100) and experienced pilots with very high (over 5000 hours) flights hours are more likely to be involved in fatal accidents.

Potter, et. al (2013), found that e-learning as part of curriculum for helicopter pilot training to be valuable in terms of safety and confidence of the pilots.

Roth (2015) recommends recurrent training sessions for each pilot. Just having the flying hours does not necessary and automatically qualify the pilots. The report concludes that the pilots not only should learn the technical abilities, they must learn how to manage fatigue and stress, communicate and be professional. The study highly recommends early investment in good basic flying skills which can be relied upon for the rest of the pilots' careers.

Makarowski et. al. (2016), indicate that higher number of flying hours will make the pilots to take unnecessary risks. This study concludes that with increased experience, the pilots seek additional conscious or unconscious stimulations. Accordingly, this leads to unnecessary risk to achieve the right level of stimulation. The so-called reasonable risk-takers are an equivalent of someone adopting the "invulnerability" attitude. This study provides an excellent overview of past studies supporting their conclusions. The study recommends ongoing examinations of scientific and psychological tests on pilots with high number of flying hours to assess their temperamental traits, aggression levels, and risk-taking tendencies.

In summary, the common theme among pilot training papers that enhance the learning include:

- One size fits all curriculum appear to be inefficient. They need to be dynamic and in some cases customized.
- Technology devices can help with better learning.
- Communication skills must be present in all trainings.
- Interactive recordings where student pilots can watch and make decisions
- High standards in terms of physical, emotional and psychological behavior for selecting candidates.
- Team training is becoming very popular among airlines.
- Recurring training for all pilots

- Inclusion of fatigue and stress in training
- Providing simulations for experienced pilots.

6. Police, Fire Fighters and First Responders

6.1 Law Enforcement Officers

Training of Police personnel in recent times has focused on development of skills, training in management and on response to acts of violence. The role of police in any geographical location is to stabilize and protect law and order (Munck, 2007), ensure the rights of citizens are protected and preservation of freedom and peaceful coexistence of members of community (Paoline and Terrill, 2007; Wright et al. 2011). Police training and development efforts in the United States began in the early 1900's (Kelling and Moore, 1988) and rapidly became more formal (Walling, 2007). The International Association of Chiefs of Police (IACP)[3] is the world's oldest and largest law enforcement body serving the needs of other law enforcement bodies across the world. It provides a number of training and development opportunities for law enforcement professionals.

August Vollmer from Berkeley was one of the pioneers to develop new techniques and methodologies for not only training but also for equipping the police officers during his time[4] (Fyfe et al., 1997). Consequently, training schools were created across the United States which focused on the use of better technology, communication, community outreach and development. Particular emphasis was laid on the conditions of self-protection and the use of force (Schauer, 2017; Frank et al., 2014; Walker, 2005; Chan, 1996). The American Society for Criminology was established in 1947 to provide further impetus to the effort. This organization currently deals with defining, measuring, controlling and creating preventive action recommendations for criminal acts.[5] Training practices developed by this organization are more focused on technical and mechanical aspects. Training still varies widely across different states, budgets and agencies (Stanilas, 2013).

[3] www.theiacp.org
[4] https://www.berkeleyside.com/2010/01/27/remembering-august-vollmer-the-berkeley-police-chief-who-created-modern-policing
[5] https://www.asc41.com/about.htm

The Academy of Criminal Justice Sciences was created in 1963 to stimulate scholarly and professional activities in criminal law and order.[6] Similarly the Police Executive Research Forum was created in 1976 to establish training and curriculum innovation research and create policy directives. Some of the curricula developed include reducing the use of police force, community oriented policing and problem solving, the use of technology such as body camera, combating violent extremism and violent crimes, cybercrimes and combating public health emergencies. For example, the Integrating Communications, Assessment and Tactics (ICAT) provides lesson plans, videos, case studies, Power Points and scenario-based training exercises to police personnel. The crisis intervention team (CIT) model is currently used extensively to help create a complete police force response to de-escalate crisis involving persons with mental illnesses. This model focusses on completing a one-time 40 hour training program which involves an understanding of different types of psychological disorders in persons and responding to criminal acts from such persons (See Figure 6). Evaluations are mostly carried out by surveys. These surveys have revealed that after going through CIT model based training, police officers are more likely to direct persons with mental illnesses to medical care rather than prisons (Bratina et al., 2018).

	Day1	Day 2	Day 3	Day 4	Day 5
8am-9am	Welcome/Introductions; Purpose & History	Alzheimer's: Issues & Community Care	Veteran Issues: PTSD; Veteran's Panel	Site visit and tour of short-term residential care facility or forensic hospital	Trauma Informed Care
9am-10am	Historical Perspectives on Treatment	Substance Abuse & Co-occurring Mental Illness	Autism Spectrum Disorder		Cop Suicide
11am-12 pm	Psychotic Disorders & Bipolar Illness: Symptoms & Medications	Cultural Diversity & Stigma of Mental Illness; Personal Perspectives	Discretion, Decision-Making and Criminal Charges	Forensic Issues	Suicide by Cop
12 pm-1 pm	Lunch	Lunch	Lunch	Lunch	Lunch
1 pm-2 pm	Skill Set: Active Listening	Depression/Anxiety: Symptoms, Medications; Risk Assessment for Suicide	Enforceable Involuntary Hospitalization (e.g., Baker Act, 2009 & Marchman Act, 1993)	Police Psychology	Hostage Negotiation Teams and CIT
2 pm-3 pm	Skill Set: Voice, Tone, Body Language, Friendliness	Risk Assessment for Violence	Children's Issues	Intellectual Disability	Excited Delirium; Personality Disorders
3 pm-4 pm	Mental Health Court Tour & Presentation; Day 1 CIT Training	Hearing Voices Audio Tapes	Resources & 211	Tour of Secure Behavioral Health Facility; Day 4 Training	Implementation, Data Collection, & State/International CIT Data
4 pm-5 pm	Evaluations	Crisis Intervention Role Play; Day 2 CIT Training Evaluations	Crisis Intervention Role Play; Day 3 CIT Training Evaluations	Evaluations	Graduation Ceremony

Figure 6: Example of a 40-hour training schedule for Crisis Intervention Team (CIT) for First Responders

One of the most cutting edge training programs is Community-Oriented Policing or COPS, which is said to be the solution to a wide variety of potential sources of conflicts and problems associated with criminal activity (Ferrandino, 2014; Rosenbaum and Lurigio, 1994). The purpose of any training program is to create standards required for successful job performance and training

[6] https://www.acjs.org/

of police and has moved a long way from a police officer getting hired and learning with another police officer on job duty (Wright et al., 2011) to incorporate the ever changing social needs of the modern day. The main theme that emerges from literature is the interdisciplinary nature of police training – involvement of psychology, exercise sciences, medicine and drugs, technology end e-commerce, social media and the application of these to problem solving particularly under stressful situations, sometimes without favorable outcomes.

Simulation and scenario based training has been useful in creating and overcoming stressful situations for police officers, per one study in California (McCraty and Atkinson, 2012). This study uses a program called Coherence Advantage Program (CAP) to help the law enforcement officers to self-regulate themselves mentally, psychologically and physically. This heightened self-awareness creates a better response to the stressful situations and increases the effectiveness of police response. Another study similarly finds that by enhancing specific adaptive responses of police officers and emergency responders to stressors, these individuals gain better control of the situation by being aware of self-response to the situation. This leads to optimal performance during times of crisis (Arnetz et al. 2009). Another study in Italian police force was undertaken to measure the effectiveness of free courses in physical efficiency (such as fitness, diet and total body conditioning) and wellness (such as meditation) found through surveys that the officers who underwent such courses felt lesser stress and consequently greater sense of wellbeing (Maran et al., 2018).

Key Lessons learned from police officer training:

- A stakeholder centric approach and focus on innovation and the use of technology to constantly deliver better results has resulted in a rapid modernization of the police force.
- Focus on addressing burning issues such as the use of excessive force, use of new technology, response to persons with mental disorders and innovative community based policing has constantly changed the landscape in training of police officers globally.
- A focus on simulating situations and making subjects self-aware of their responses under stressful situations improves the effectiveness of police officer response to such situations.

6.2 Firefighters and Emergency Responders

First responders are constantly faced with stressors and incidents that cause long term negative health outcomes because of their diverse, dangerous and potentially life-threatening job responsibilities. To overcome their day-to-day challenges, firefighters have to complete the National Fire Protection Association's (NFPA) Firefighter I and II courses. Although, according to NFPA itself, its professional standards are neither required nor intended as a formal training set of documents. However, many training institutions use the NFPA standards. Further, two accreditation agencies Pro Board and IFSAC require training institutions to meet or exceed the standards set by the NFPA.[7] It was found that on an average, the firefighter training curriculum involves approx. 270 hours of theoretical and skill-based training. This training includes fire awareness, firefighter I and II, HAZMAT awareness and operations and emergency driver training.

Firefighters are constantly exposed to dangerous situations, hazardous chemicals, life threatening scenarios and traumatic accidents, sometimes with insufficient and inadequate equipment. They often face long shifts, sometimes over weekends away from their friends and family. It is believed that this job-related stress causes greater than normal rates of suicidal ideas and attempts among firefighters and emergency medical service personnel (Martin et al., 2017). Even otherwise these personnel report incidents of alcohol dependence, depression and posttraumatic stress disorder (PTSD) severity (Stanley et al., 2016).

Firefighters, law enforcement officers and first medical responders also have to deal with other stressors such as threats of violence, assaults and fatalities (Regambal et al., 2015; Regehr et al., 2003a and Regehr et al., 2003b). This incidence of stress causing agents at work has also been observed in the case physicians, medical trainees and nurses (Regehr et al., 2014, West et al., 2016). In the most recent literature, psychological resilience has been shown to be variable and learnable process of creating a way to overcome stress over short run as well as the long run (Jensen and Fraser, 2016; Schnyder, 2014; Marks et al., 2017).

There is some evidence that psychological resilience-which is a learnable technique, to cope with mechanisms and incidents that create burnout in occupations involving severe or chronic

[7] www.nfpa.org (research foundation reports) accessed September 4, 2018

stress. Consequently, a new technique- Mindfulness-Based Resilience Training (MBRT) is an intervention technique which has been developed to gradually build this resilience and help the subjects such as first responders to overcome negative health outcomes such as burnout due to severe job stress (Kaplan et al., 2017). It is an 8-week course which involves discussion, learning, body scans, and experiential exercises. The content has been modified to suit the needs of first responders such as critical incidents, public scrutiny and job satisfaction, or the lack thereof.

There are surveys used to measure mindfulness, burnout and the connections between these two and resilience. One of these surveys is the Five Facet Mindfulness Questionnaire (FFMQ; Williams et al., 2014; Baer et al., 2006), which is commonly used to measure the tendency of individuals to pay attention to the psychological needs of the subjects. Another useful survey which measures resilience is the Brief Resilience Scale (BRS; Smith et al., 2008). Finally, the Maslach Burnout Inventory (MBI; Maslach et al., 2001) and Oldenburg Burnout Inventory (OBI; Demerouti et al., 2010) are some ways to measure the level of mental and emotional breakdown of a person as a result of a stressful situation. This link between stress, resilience and burnout are explored by Hao et al. (2015) in the Civil Servants of China. The link between emotional exhaustion and a lack of quality care has been studied by Garcia and Calvo (2012) in studying the exhaustion of nursing staff of Spain. Another study on employees of a public hospital in Brazil found a connection between resilience and decreased incidence of burnout (Carneiro et al., 2018). Kaplan et al. (2017) report that increased mindfulness was linked with increased resilience, which was in turn linked with decreased burnout. However, the same study finds that resilience and burnout are linked with each other independently (from mindfulness).

Behavioral Health Training Program (BHT) is another training system that has been created with an intention of providing education and awareness to firefighters created by the Broward Sherriff's Office (BSO) for fire rescue and emergency services personnel (Steinkopf et al., 2016). The BSO partnered with academicians from the College of Psychology from Nova Southeastern University and developed modules on overview, stress, depression, sleep, substance abuse and suicide. The understanding of participants was measured using a questionnaire and it was observed that not only did the understanding of participants go up, but the participants also appreciated the fact that a training module like this improves their self-awareness as well as stress coping mechanisms.

Simulation involves the dynamic generation of real world fire related phenomena observed in the past using controlled environment either in an actual experimental set-up or using a computer based environment. The most innovative research and training for firefighters being developed globally is by using simulations and virtual reality. Part of the problem with training with practical exercises is that these are important to acquaint the trainees with new equipment and some application of fire under controlled conditions. However, in such situations fire does not behave in the same way as in a real situation. Furthermore, simulations allow trainers to reproduce complex scenarios such as forest fires and calamitous fire incidents in urban areas, which cannot be produced in real life until the trainees encounter such situations firsthand. Consequently, the fire knowledge is left more to anecdotal training and fire behavior is left at a theoretical level until the trainees encounter it firsthand (Moreno et al., 2014). Virtual reality based simulators enable the recreation of such scenarios and allow the trainees to experience these firsthand even before the actual incidents are witnessed in real life. One such simulator has been studied by Moreno et al. (2014). This incorporates fire spread algorithms for both forest and city scenarios and help users learn extinguishment under a variety of climate conditions such as wind, rain, temperature and humidity.

Web based interactive game like training has also been developed to improve firefighter learning skills at various training sessions. One such system is Advanced Learning in Integrated Visual Environments (ALIVE) published in their study by Wener et al. (2015). ALIVE training is a product of collaboration between the Fire Dept. of New York (FDNY) and New York University Polytechnic School of Engineering (NYU-Poly) and the National Institute of Standards and Technology (NIST). It uses audio, video and graphics with its modules to present real world fire fighting scenarios and asks the users to formulate and present a plan to deal with such scenarios. These scenarios include cases such as fire in high-rises, modern light-weight houses with layouts different from conventional layouts such as open floor plans and also sleep related issues such as

fatigue from long shifts. There are several vendors who also provide simulation based training such as Simshare[8], VSTEP[9], FLAME-SIM[10] by firerescue1, xvrsim[11] and digitalcombustion[12].

Cha et al. (2012) propose the use of a fire training simulator from the beginning of a fire to demonstrate liquid and gas flows at different temperatures. The scenarios differ by fire sizes, air flow, avatar control and starting point (tunnel entrance/ exit/ shelter etc.). The scenarios include overview, training and experience of fire. The study uses a head mounted display instead of a device screen interface to make the experience more immersive.

Another multimedia based firefighting tool developed by Kolmanic et al. (2013) for both full-time and volunteer firefighter trainees in Slovenia. The tool called SIN, is based on the photographic images and videos of environments familiar to the firefighters with incidents involving both sound and graphics. The situations that can be used for training purposes have been extended to traumatic accidents and involvement of hazardous chemicals. While a number of studies focus on the scenario, firefighters need to be aware of the changes occurring around themselves including their gear. One such study focusses on the simulation of fire accidents in the petrochemical industry to provide a visual and sensory experience for firefighter training. The study measures the change in temperature and humidity and creates this in a specialized attire worn by trainee which regulates temperature using heating units and coolant fans embedded in the cloth. Another study by Horn et al. (2015) studies the physiological response of firefighters in three different scenarios during training. These are environment chamber simulated fire, treadmill activities, simulated burning building. The paper discusses the physiological responses of the subjects in each of these scenarios, such as heart rate, core temperature and the level of mental stress (also studied by Ensari et al., 2016) and finds that the closed environment of chamber simulated fire produces the most stressful response from firefighters implying greater exposure to such training.

---

[8] http://simsushare.com (accessed 9-22-2018)
[9] https://www.vstepsimulation.com (accessed 9-22-2018)
[10] https://www.firerescue1.com/company-directory/FLAME-SIM (accessed 9-22-2018)
[11] http://www.xvrsim.com (accessed 9-22-2018)
[12] http://www.digitalcombustion.com/about-digital-combustion (accessed 9-22-2018)

Simulation based training has been used in Norwegian Emergency Medical Services to promote situation awareness, decision making and communication, leadership, managing and coping with stress and fatigue and teamwork. A study by Abrahamsen et al. (2015) finds using survey based questionnaire that the flight operators have had greater access to simulation based training than the health enterprise employees. Further, simulation based training enhances non-technical training skills for all employees.

Implications for ATC and ATSS: the forefront of research on training of first responders has lately been focusing on mental stress, work hours, fatigue and burnout and strategies to overcome these. In addition to conventional training modules, these subjects are now undergoing training to understand the stressful situations that they can potentially face through increased access to simulation based training using augmented and virtual reality. Some countries are also trying immersive virtual reality by recreating the accident conditions. Studies have found greater level of comfort from first responders after simulation based training. Similarly, ATC should be exposed to simulations of stressful situations recreated from past experiences to be exposed to such situations. The other implication for ATC is to reduce stress by understanding how to cope with it using mindfulness training modules.

In summary, the common theme in training of first responders and law enforcement officers include:

- Training of fundamental principles in a classroom based environment.
- Frequent feedback, examination and accreditation (where applicable) of training modules
- Focus on physical, mental and interdisciplinary topics which include
    - Physical activities
    - Mental stress evaluation and coping techniques
    - Mindfulness training and exposure to potential stressors
- Use of AR and VR based systems

7. Some additional lessons learned from other industries

7.1 Military Training:

Wampler, et. al (2006) published a comprehensive report which summarizes lessons learned from military training over a six year period. This study was initiated as the US army was engaged in two wars in 2006. The army was interested in identifying ways to train the soldiers and combat personnel to meet and optimize their strategic goals and resources. The research included training efforts for different learning materials, new equipment, digital systems, and a variety of experimental programs. The authors reviewed and analyzed the training lessons from 1998-2004. Their research concluded the followings to have the highest impact on positive soldier/ combat personnel trainings:

- Interactive courseware
- Correct sequence of learning materials
- Use of technology
- Selection process
- Simplifying learning materials
- Allocating sufficient training time.

7.2 Computer Based Training

Bedwell and Salas (2010), review computer based training (CBT) for diverse industries for providing systematic and structural learning. Their study offers 10 practical considerations on successful CBT design and development and how to avoid failures. Some of these considerations include:

- Emphasis on content design and effectively match delivery with learning outcomes.
- CBT not the only answer to all trainings and development needs
- Blended learning including face-to-face training, discussions and live scenarios through CBT
- Thoroughly define training purpose
- CBT needs to be integrated with metrics to evaluate the effectiveness.

- Rely on principles of multimedia design that works.
- Instructor is instrumental to the process
- CBT needs to be adaptive to individuals as learning doffers among different people.

7.3 Labor Market trainings

Kletzer and Koch (2004), conduct a study on identifying what factors impact labor skills building at a macro level among several member countries in Organization for Economic Co-operation and Development (OECD) including USA. They analyze factors such as percentages of expenditures on labor training as percentage of GDP, unemployment rates, percentages of adult-aged education and financial support offered by private and public sectors. Their main conclusion related to skill building include:

- Workers increasingly need problem-solving skills
- Skills that cross jobs as teamwork and quality control replace simpler and more hierarchical production processes.

7.4 African Famers Training

De Rosa, et. al. (2016), review the training practices for farmers in Africa developed by The Food and Agriculture Organization of the United Nations (FAO). They identify four initiatives that have proved to be very successful for training.

- Rigorous selection of participants
- Effective communications
- Integration of learning into strategic engagement
- Meaningful Assessments

7.5 Nuclear Power Plants

Lin et al. (2014) discuss and develop EID in advanced main control rooms (in a nuclear power plant). This study measures the interactions of manager-operator (MO), reactor-operator (RO) and assistant reactor operator (ARO). The study concludes that assistant reactor operators showed good understanding in skill-based level behavior while manager-operators showed good understanding in knowledge-based level behavior. The key finding from this paper is that the use of role play in simulated scenario to measure the interactions of MO-RO-AROs chosen randomly from sample of individuals working at the plant.

Lee et al. (2015) measure the situation awareness (SA) of operators in the dynamic main control room (MCR) of a nuclear power plant (npp). They define tasks required for each level of understanding and propose to measure the situation awareness (SA) of the operators under these conditions. SA1 corresponds to skill level understanding, SA2 to rule level and SA3 to knowledge level understanding. The authors observe the team level SA and base their finding on the assumption that team communication is a direct indicator of team SA. These are either through observation or direct participation. Key finding from the paper is that the measure scores of SA by providing a hypothetical scenario to random teams from 9 different nuclear reactors in a high fidelity full-scope simulator reactor.

Overall, ecological interface design helps create familiarity with complex systems such as power grids and nuclear power plants. Situational awareness helps trainees tackle sudden, unexpected and complex problems with relatively more ease.

8. Recommendations, Best Training Practices and Concluding Remarks

In each area that we studied, there is traditional, standard training methodology, which is often lecture based and includes some kind of assessment activity at the end of the program. All these fields include practical, hands -on training with professionals who train the trainees and include an exposure of trainees to case based scenarios. Training modules often follow a scaffolding process and this process often follows a diverse (and sometimes complicated) path to conclusion. Further, there is diversity across geographical locations and across professions. However, all these professions provide ample practical training before the trainee is inducted formally as a team member/ professional. On-job training often follows with the trainees being at an advanced level of understanding of the demands and complications of their job.

Some of the best practices that emerged as a trend across these disciplines are the use of simulations, enactment of real life situations in training environment and prompt assessment and feedback with a subsequent follow-up. Two more trends arising out of our research are the focus on new technology and on the levels of stress that the job demands from the professionals. We observe this across many different geographical locations (countries) of the world and not just in the countries of the western world.

The evolution of modern day training practices focus on the enhanced lecture classes with a planned and thought out curriculum, that is divided in small achievable outcomes. These classes employ the use of technology, such as lectures, cases and situation-studies being available on video on demand to the trainees. Also there is a noticeable shift toward careful planning and periodic revisit of the curriculum from independent regulatory bodies (such as the NFPA, Pro Board and IFSAC in the case of firefighter training). We also observe that there is a focus on the learning behavior and psychological aspects of development in trainees (i.e. the ability to handle stress), especially because of constantly growing demands from these professionals through long, uncertain shifts and with the expectation of sudden and sometimes extremely stressful events.

From nursing, we learn the key practices of standardization in curriculum, whether it is patients or the type of activity/ situation, to the development of psychomotor skills, laboratory training, use of computer based programs for training and assessment , feedback and follow-up.

A surgeon carries a lot of responsibilities: those of patients' health and well-being, of the reputation of their own and of the hospital or the institution they represent. A surgeon is a leader, who carries out several complex tasks. Therefore, surgeon training involves several stages, including but not limited to theoretical lessons, practical training, acclimatization of the operating room, learning the technique, trying the technique and performing surgery under supervision. Furthermore, the focus of training and planning of different training modules varies widely across different fields of surgery, because of varied levels of complexity.

Surgeon training initially focuses on the skill inventory that the trainee possesses before the training and then is adapted to suit the trainee's needs. It also focusses on understanding the behavioral aspects of patients and how training can systematically address the gaps in skill, rule and knowledge levels. However, once a surgeon trainee passes the classroom level, training becomes more sophisticated with a focus on these gaps. That is where simulation based-training plays a part. Numerous reviews on surgeon training have shown simulations with at least medium fidelity and medium level of immersion to be optimal for skill development and practice.

The cost factor remains a major obstacle in implementing high fidelity high immersion based simulations. Further, due to regulatory and ethical considerations, training of surgery is moving away from human or animal subjects. This void is being filled by VR-based surgery training. Measurement of performance is carried out qualitatively and quantitatively through surveys, tests and feedback from trainers. Feedback can be provided after each stage and/ or at the end of the complete module. Further, there is evidence that reinforcement of training, by revisiting what was learned in the past.

The role of pilots is extremely critical in an every growing world of aviation. According to a survey by IATA[13], the total number of passengers travelling by air globally has increased from 2.7 billion in 2010 to 3.8 billion in 2016. Similarly the number of flights globally increased from 27.8 million in 2010 to 35.4 million in 2016. One of the key lessons learned from pilot training are that the training curricula should be flexible, depending on ability and outcome desired from trainers and students. Team based training should receive equal importance as individual training.

[13] https://www.iata.org/publications/economics/Reports/Industry-Econ-Performance/Central-forecast-mid-year-2018-tables-v1.0.pdf Accessed Sept. 23, 2018

Training should focus on preparing checklists, which are known to increase safety. Discussions should focus more on scenario based analysis, which has been shown to promote critical thinking among pilots. Similar to surgery, training should focus on the recurrence of training and revalidation of qualification. Stress should be laid on the development of interpersonal skills. Use of active learning strategies and the comparison of pre- vs post-training assessment scores should be made to evaluate the efficacy of training. Pilot training also stresses on the importance of handling unexpected, potentially stressful and life threatening situations.

First responders and law enforcement officers are constantly under stress. In addition to responding to emergency, they also have to perform routine tasks. Therefore, there are lessons to be learnt for both ATC and ATSS. Both firefighters and police officers undergo classroom training, similar to the other professionals discussed previously in this section. In addition, firefighters and law enforcement officers also have to undergo physical training and evaluation. However, the key aspect of focus have shifted over to simulation of stressful situations either through case based training or through the use of AR and VR. Additionally, law enforcement is increasingly focusing on the involvement of the community. The other focus of training in both these areas is on understanding of and coping with stress. Consequently, some time is being spent on self-awareness and mindfulness. Development in technology and gamification has also contributed to better ways of dealing with stressful situations.

In this study, we also studied training in nuclear power plants, agriculture, military and labor markets . The key lessons to be learnt are the increased use of and focus on technology, not only on improving the way the job is done, but also the way training is carried out. This is in line with all other fields that we have studied and presented in this paper. Whether it is gamification or the use of interactive audio-video enabled hands-on skill development with prompt feedback or the use of training for the purpose of communication and team-based learning strategies. Situation awareness is similarly, a very important factor and can be promoted by hypothetical scenarios, roleplays and high-fidelity simulations

9. References

ADMS-Fire, *http://www.trainingfordisastermanagement.com/*

Aguilar-Moya, R., Melero-Fuentes, D., Navarro-Molina, C., Aleixandre-Benavent, R., & Valderrama-Zurián, J. C. (2014). Disciplines and thematics of scientific research in police training (1988-2012). *Policing: An International Journal of Police Strategies & Management*, *37*(4), 696-711.

Ahlberg, G., Enochsson, L., Gallagher, A. G., Hedman, L., Hogman, C., McClusky III, D. A., Ramel, S., Smith, C. D. & Arvidsson, D. (2007). Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *The American journal of surgery*, *193*(6), 797-804.

Alcañiz, M., Rey, B., Tembl, J., & Parkhutik, V. (2009). A neuroscience approach to virtual reality experience using transcranial Doppler monitoring. *Presence: Teleoperators and Virtual Environments*, *18*(2), 97-111.

Alessi, S. M. (1988). Fidelity in the design of instructional simulations. *Journal of computer-based instruction*.

Alinier, G., Hunt, B., Gordon, R., & Harwood, C. (2006). "Effectiveness of intermediate-fidelity simulation training technology in undergraduate nursing education." *Journal of advanced nursing*, 54(3), 359-369.

Alpert, G. P., & Dunham, R. G. (1988). *Policing urban america*. Prospect Heights, IL: Waveland Press.

America's Army, https://www.americasarmy.com/

Amin, A., Whiteley, I., & Shearer, R. (2016). Walker K. iViewExpert: using a video-cued recall and debrief method from astronaut, pilot and firefighter training, to understand and transfer technical medical expertise. *Rural and Remote Health*, *16*, 4111.

André Pinheiro, Paulo Fernandes, Ana Maia, Gonçalo Cruz, Daniela Pedrosa, Benjamim Fonseca, Hugo Paredes, Paulo Martins, Leonel Morgado, Jorge Rafael (2014), Development of a mechanical maintenance training simulator in OpenSimulator for F-16 aircraft engines, *Entertainment Computing*, Volume 5, Issue 4, 2014, Pages 347-355

Arnetz, B. B., Nevedal, D. C., Lumley, M. A., Backman, L., & Lublin, A. (2009). Trauma resilience training for police: Psychophysiological and performance effects. *Journal of Police and Criminal Psychology*, *24*(1), 1-9.

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, *13*(1), 27-45.

Bagai, A., O'Brien, S., Hatim Al Lawati, M. D., Goyal, P., Ball, W., Grantcharov, T., & Fam, N. (2012). Mentored Simulation Training Improves Procedural Skills in Cardiac Catheterization: a randomized, controlled pilot study. *Circulation Cardiovascular Interventions*, *5*, 672-679.

Banks, E. H., Chudnoff, S., Karmin, I., Wang, C., & Pardanani, S. (2007). Does a surgical simulator improve resident operative performance of laparoscopic tubal ligation?. *American journal of obstetrics and gynecology*, *197*(5), 541-e1.

Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. *Advances in Human Factors Ergonomics*, *19*, 699-699.

Baumgartner, T., Valko, L., Esslen, M., & Jäncke, L. (2006). Neural correlate of spatial presence in an arousing and noninteractive virtual reality: an EEG and psychophysiology study. *CyberPsychology & Behavior*, *9*(1), 30-45.

Beaubien, J. M., & Baker, D. P. (2004). The use of simulation for training teamwork skills in health care: how low can you go?. *BMJ Quality & Safety*, *13*(suppl 1), i51-i56.

Bedwell, W. L., & Salas, E., (2010). Computer-based training: Capitalizing on lessons learned, International Journal of Training and Development, Vol. 14, Issue: 3, pp 239-249.

Belyea, D. A., Brown, S. E., & Rajjoub, L. Z. (2011). Influence of surgery simulator training on ophthalmology resident phacoemulsification performance. *Journal of Cataract & Refractive Surgery*, *37*(10), 1756-1761.

Beyer, L., De Troyer, J., Mancini, J., Bladou, F., Berdah, S. V., & Karsenty, G. (2011). Impact of laparoscopy simulator training on the technical skills of future surgeons in the operating room: a prospective study. *The American Journal of Surgery*, *202*(3), 265-272.

Biglino, Giovanni, Claudio Capelli, Despina Koniordou, Di Robertshaw, Lindsay-Kay Leaver, Silvia Schievano, Andrew M. Taylor, and Jo Wray. "Use of 3D models of congenital heart disease as an education tool for cardiac nurses." *Congenital heart disease*, 12(1) (2017): 113-118.

Bliss, Maria, and Leanne M. Aitken. "Does simulation enhance nurses' ability to assess deteriorating patients?." *Nurse education in practice*, 28 (2018): 20-26.

Booth, Richard G., Christina Ko Scerbo, Barbara Sinclair, Michele Hancock, David Reid, and Eileen Denomy. "Exploring learning content and knowledge transfer in baccalaureate nursing students using a hybrid mental health practice experience." *Nurse education today*, 51 (2017): 57-62.

Bratina, M. P., Carrero, K. M., Kim, B., & Merlo, A. V. (2018). Crisis intervention team training: when police encounter persons with mental illness. *Police Practice and Research*, 1-18.

Burke, E., Hobson, C., & Linsky, C., (1997). Large Sample Validations of Three General Predictors of Pilot Training Success. *The International Journal of Aviation Psychology*, 7:3, 225-234.

Caird, J. K. (1996). Persistent issues in the application of virtual environment systems to training. In *Human Interaction with Complex Systems, 1996. HICS'96. Proceedings., Third Annual Symposium on* (pp. 124-132). IEEE.

Campbell, J. S., Castaneda, M., & Pulos, S., (2009). Meta-Analysis of Personality Assessments as Predictors of Military Aviation Training Success. *The International Journal of Aviation Psychology*, 20:1, 92-109

Cant, Robyn P., and Simon J. Cooper. "Use of simulation-based learning in undergraduate nurse education: An umbrella systematic review." *Nurse education today*, 49 (2017): 63-71.

Carneiro, É. M., Navinchandra, S. A., Vento, L., Timóteo, R. P., & de Fátima Borges, M. (2018). Religiousness/Spirituality, Resilience and Burnout in Employees of a Public Hospital in Brazil. *Journal of religion and health*, 1-9.

Caro, P. W., (1973). Aircraft Simulators and Pilot Training. *Human Factors*. 15, 502–509.

Chan, J. (1996). Changing police culture. *The British Journal of Criminology*, *36*(1), 109-134.

Chang, Y. H., Yang, H. H., & Hsiao, Y. J. (2016). Human risk factors associated with pilots in runway excursions. *Accident Analysis and Prevention*, 227-237.

Clemente, M., Rodríguez, A., Rey, B., & Alcañiz, M. (2014). Assessment of the influence of navigation control and screen size on the sense of presence in virtual reality using EEG. *Expert Systems with Applications*, *41*(4), 1584-1592.

Cornelissen, F., Mark A Neerincx, Nanja Smets, Leo Breebaart, Paul Dujardin, Mikael Wolff, “Gamification for Astronaut Training,” *SpaceOps 2012 Conference*, (2012), Stockholm, Sweden

Cosman, P., Hugh, T., Shearer, C., Merrett, N., Biankin, A., & Cartmill, J. A. (2007). Skills acquired on virtual reality laparoscopic simulators transfer into the operating room in a blinded, randomised, controlled trial. *Stud Health Technol Inform*, *125*(1), 76-81.

Crochet, P., Aggarwal, R., Knight, S., Berdah, S., Boubli, L., & Agostini, A. (2017). Development of an evidence-based training program for laparoscopic hysterectomy on a virtual reality simulator. *Surgical endoscopy*, *31*(6), 2474-2482.

Dankelman, J., Wentink, M., Grimbergen, C. A., Stassen, H. G., & Reekers, J. (2004). Does virtual reality training make sense in interventional radiology? Training skill-, rule-and knowledge-based behavior. *Cardiovascular and interventional radiology*, *27*(5), 417-421.

Dawe, S. R., Pena, G. N., Windsor, J. A., Broeders, J. A. J. L., Cregan, P. C., Hewett, P. J., & Maddern, G. J. (2014). Systematic review of skills transfer after surgical simulation-based training. *British Journal of Surgery*, *101*(9), 1063-1076.

Demerouti, E., Mostert, K., & Bakker, A. B. (2010). Burnout and work engagement: a thorough investigation of the independency of both constructs. *Journal of occupational health psychology*, *15*(3), 209.

Deutschmann, M. W., Yunker, W. K., Cho, J. J., Andreassen, M., Beveridge, S., & Bosch, J. D. (2013). Use of a Low-fidelity simulator to improve trans-nasal fibre-optic flexible laryngoscopy in the clinical setting: a randomized, single-blinded, prospective study. *Journal of Otolaryngology-Head & Neck Surgery*, *42*(1), 35.

Drivalou, S., & Marmaras, N. (2009). Supporting skill-, rule-, and knowledge-based behaviour through an ecological interface: An industry-scale application. *International Journal of Industrial Ergonomics*, *39*(6), 947-965.

Earthlight, *https://www.earthlightvr.com/*

Edge, *https://www.dhs.gov/science-and-technology/EDGE*

El-beheiry, M., Mccreery, G., & Schlachta, C. M. (2017). A serious game skills competition increases voluntary usage and proficiency of a virtual reality laparoscopic simulator during first-year surgical residents' simulation curriculum. *Surgical Endoscopy*, *31*(4), 1643-1650.

Ende, A., Zopf, Y., Konturek, P., Naegel, A., Hahn, E. G., Matthes, K., & Maiss, J. (2012). Strategies for training in diagnostic upper endoscopy: a prospective, randomized trial. *Gastrointestinal endoscopy*, *75*(2), 254-260.

Ensari, I., Motl, R. W., Klaren, R. E., Fernhall, B., Smith, D. L., & Horn, G. P. (2017). Firefighter exercise protocols conducted in an environmental chamber: developing a laboratory-based simulated firefighting protocol. *Ergonomics*, *60*(5), 657-668.

Ferdinand Cornelissen, Mark A Neerincx, Nanja Smets, Leo Breebaart, Paul Dujardin, Mikael Wolff (2012), Gamification for Astronaut Training, *SpaceOps 2012 Conference*, Stockholm, Sweden

Ferlitsch, A., Schoefl, R., Puespoek, A., Miehsler, W., Schoeniger-Hekele, M., Hofer, H., Gangi, A. & Homoncik, M. (2010). Effect of virtual endoscopy simulator training on performance of upper gastrointestinal endoscopy in patients: a randomized controlled trial. *Endoscopy*, *42*(12), 1049-1056.

Ferrandino, J. A. (2014). An integrated theory for the practical application of "governance-based policing". *Policing: An International Journal of Police Strategies & Management*, *37*(1), 52-69.

Foss, B., K. Solbjørg, Ole, Ivar Eikaas, Tor and Frank, Jakobsen, " Game play in vocational training and engineering education," *Modeling, Identification and Control*, (2007), 28. 10.4173/mic.2007.2.1.

Flamtrainer, *http://flaimtrainer.com/*

Frank, K.,IV, Frank, J., & Liederbach, J. (2014). Understanding police use of force. *Policing*, *37*(3), 558-578.

Franzeck, F. M., Rosenthal, R., Muller, M. K., Nocito, A., Wittich, F., Maurus, C., Dindo, D., Clavien, P.A. & Hahnloser, D. (2012). Prospective randomized controlled trial of simulator-based versus traditional in-surgery laparoscopic camera navigation training. *Surgical endoscopy*, *26*(1), 235-241.

Fried, M. P., Kaye, R. J., Gibber, M. J., Jackman, A. H., Paskhover, B. P., Sadoughi, B., Schiff, B., Fraioli, R. E. & Jacobs, J. B. (2012). Criterion-based (proficiency) training to improve surgical performance. *Archives of Otolaryngology–Head & Neck Surgery*, *138*(11), 1024-1029.

Fried, M. P., Sadoughi, B., Gibber, M. J., Jacobs, J. B., Lebowitz, R. A., Ross, D. A., Bent III, J. P., Parikh, S. R., Sasaki, C. T. & Schaefer, S. D. (2010). From virtual reality to the operating room: the endoscopic sinus surgery simulator experiment. *Otolaryngology—Head and Neck Surgery*, *142*(2), 202-207.

Gala, R., Orejuela, F., Gerten, K., Lockrow, E., Kilpatrick, C., Chohan, L., Green, C., Vaught, J., Goldberg, A. & Schaffer, J. (2013). Effect of validated skills simulation on operating room performance in obstetrics and gynecology residents: a randomized controlled trial. *Obstetrics & Gynecology*, *121*(3), 578-584.

Gallagher, A. G., Henn, P. J., Neary, P. C., Senagore, A. J., Marcello, P. W., Bunting, B. P., Seymour, N. E., & Satava, R. M. (2018). Outlier experienced surgeon's performances impact on benchmark for technical surgical skills training. *ANZ journal of surgery*, *88*(5), E412-E417.

Gallagher, A. G., Ritter, E. M., Champion, H., Higgins, G., Fried, M. P., Moses, G., Smith, C. D. & Satava, R. M. (2005). Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Annals of surgery*, *241*(2), 364.

Gharaibeh, Besher, Issa Hweidi, and Ahmed Al-Smadi. "Attitudes and perception of baccalaureate nursing students toward educational simulation." *Cogent Education*, 4.1 (2017): 1360063.

Gottlieb, R., Baechle, M. A., Janus, C., & Lanning, S. K. (2017). Predicting performance in technical preclinical dental courses using advanced simulation. *Journal of dental education*, *81*(1), 101-109.

Gottlieb, R., Lanning, S. K., Gunsolley, J. C., & Buchanan, J. A. (2011). Faculty impressions of dental students' performance with and without virtual reality simulation. *Journal of dental education*, *75*(11), 1443-1451.

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clinical psychology review*, *37*, 1-12.

Hales, B. M., & Pronovost, P. J. (2006). The checklist - a tool for error management and performance improvement. *Journal of Critical Care*, 231-235.

Hamari, J., Jonna Koivisto, Harri Sarsa, “Does Gamification Work? – a literature Review of Empirical Studies on Gamification”, *47th Hawaii International Conference on System Science*, (2014)

Hamman, W. R. (2004). The complexity of team training: What we have learned from aviation and its applications to medicine. *Quality & Safety in Health Care, 13*

Hao, S., Hong, W., Xu, H., Zhou, L., & Xie, Z. (2015). Relationship between resilience, stress and burnout among civil servants in Beijing, China: Mediating and moderating effect analysis. *Personality and Individual Differences*, *83*, 65-71.

Haycock, A., Koch, A. D., Familiari, P., van Delft, F., Dekker, E., Petruzziello, L., Haringsma, J. & Thomas-Gibson, S. (2010). Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training. *Gastrointestinal endoscopy*, *71*(2), 298-307.

Hogle, N. J., Chang, L., Strong, V. E. M., Welcome, A. O. U., Sinaan, M., Bailey, R., & Fowler, D. L. (2009). Validation of laparoscopic surgical skills training outside the operating room: a long road. *Surgical endoscopy*, *23*(7), 1476-1482.

Hoglund, P., “Gamification in Training: engagement motivation”, thesis, *Hanken School of Economics*, Helsinki, (2014)

Hong, S. J., Lee, K. S., Seol, E. S., & Young, S. (2016). Safety perceptions of training pilots based on training institution and experience. *Journal of Air Transport Management*, 213-221.

Horn, G. P., Kesler, R. M., Motl, R. W., Hsiao-Wecksler, E. T., Klaren, R. E., Ensari, I., M. N., Fernhall, B., & Rosengren, K. S. (2015). Physiological responses to simulated firefighter exercise protocols in varying environments. *Ergonomics*, *58*(6), 1012-1021.

Howells, N. R., Gill, H. S., Carr, A. J., Price, A. J., & Rees, J. L. (2008). Transferring simulated arthroscopic skills to the operating theatre: a randomised blinded study. *The Journal of bone and joint surgery. British volume*, *90*(4), 494-499.

Hseino, H., Nugent, E., Lee, M. J., Hill, A. D., Neary, P., Tierney, S., Moneley, D. & Given, M. (2012). Skills transfer after proficiency-based simulation training in superficial femoral artery angioplasty. *Simulation in Healthcare*, *7*(5), 274-281.

Huber, T., Paschold, M., Hansen, C., Wunderling, T., Lang, H., & Kneist, W. (2017). New dimensions in surgical training: immersive virtual reality laparoscopic simulation exhilarates surgical staff. *Surgical endoscopy*, *31*(11), 4472-4477.

IFALPA Pilot Training Standards: Guide for Best Practices. (2012). *The International Federation of Air Line Pilots' Associations*.

Imber, S., Shapira, G., Gordon, M., Judes, H., & Metzger, Z. (2003). A virtual reality dental simulator predicts performance in an operative dentistry manikin course. *European Journal of Dental Education*, *7*(4), 160-163.

Jeenal Vora, Santosh Nair, Anand K Gramopadhye, Andrew T Duchowski, Brian J Melloy, Barbara Kanki (2002), Using virtual reality technology for aircraft visual inspection training: presence and comparison studies, *Applied Ergonomics*, Volume 33, Issue 6, 2002, Pages 559-570.

Jerome Planchon, Anthony Vacher, Jeremy Comblet, Eric Rabatel, Françoise Darses, Alexandre Mignon, Pierre Pasquier (2018), Serious game training improves performance in combat life-saving interventions, *Injury*, Volume 49, Issue 1, 2018, Pages 86-92,

Johnston, Sandra, Fiona Coyer, and Robyn Nash. "Simulation debriefing based on principles of transfer of learning: A pilot study." *Nurse education in practice*, 26 (2017): 102-108.

Källström, R., Hjertberg, H., & Svanvik, J. (2010). Impact of virtual reality-simulated training on urology residents' performance of transurethral resection of the prostate. *Journal of endourology*, *24*(9), 1521-1528.

Katie Li, Mark Hall, Pablo Bermell-Garcia, Jeffrey Alcock, Ashutosh Tiwari, Mar González-Franco (2017), Measuring the Learning Effectiveness of Serious Gaming for Training of Complex Manufacturing Tasks, *Simulation & Gaming*, Volume: 48 issue: 6, page(s): 770-790, December 2017

Kaplan, J. B., Bergman, A. L., Christopher, M., Bowen, S., & Hunsinger, M. (2017). Role of resilience in mindfulness training for first responders. *Mindfulness*, *8*(5), 1373-1380.

Kerfoot, B.P., N. Kissane, "The Use of Gamification to Boost Residents' Engagement in Simulation Training," *JAMA Surg. 2014*, 149(11), 1208–1209.

Kerfoot BP, Kissane N (2014), The Use of Gamification to Boost Residents' Engagement in Simulation Training. *JAMA Surgery*, Vol 149, No. 11, pages 1208–1209

Kletzer, L. G., & Koch, W. L., (2004). International Experience with Job Training: Lessons for the United States. Kalamazoo, MI: W. E. Upjohn Institute, pp 245-288.

Kober, S. E., & Neuper, C. (2012). Using auditory event-related EEG potentials to assess presence in virtual reality. *International Journal of Human-Computer Studies*, *70*(9), 577-587.

Kober, S. E., Kurzmann, J., & Neuper, C. (2012). Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: an EEG study. *International Journal of Psychophysiology*, *83*(3), 365-374.

Koglbauer, I. (2015). Simulator training improves the estimation of collision parameters and the performance of student pilots. *Procedia - Social and Behavioral Sciences*, 261-267.

Langridge, Benjamin, Sheikh Momin, Ben Coumbe, Evelina Woin, Michelle Griffin, and Peter Butler. "Systematic review of the use of 3-dimensional printing in surgical teaching and assessment." *Journal of surgical education*, 75(1), (2018): 209-221.

Larsen, C. R., Soerensen, J. L., Grantcharov, T. P., Dalsgaard, T., Schouenborg, L., Ottosen, C., Schroeder, T. V. & Ottesen, B. S. (2009). Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *Bmj*, *338*, b1802.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments*, *10*(3), 282-297.

Lin, C. J., Shiang, W. J., Chuang, C. Y., & Liou, J. L. (2014). Applying the skill-rule-knowledge framework to understanding operators' behaviors and workload in advanced main control rooms. *Nuclear Engineering and Design*, *270*, 176-184.

Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human factors*, *30*(5), 583-598.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms?. *Cognitive Psychology*, *22*(1), 1-35.

LUDUS, *http://www.ludus-vr.com/en/portfolio/firefighter-services/*

Makarowski, R., Makarowski, P., Smolicz, T., & Plopa, M. (2016). Risk profiling of airline pilots: Experience, temperamental traits and aggression. *Journal of Air Transport Management*, 298-305.

Mancini, A. D., Littleton, H. L., & Grills, A. E. (2016). Can people benefit from acute stress? Social support, psychological improvement, and resilience after the Virginia Tech campus shootings. *Clinical Psychological Science*, *4*(3), 401-417.

Maran, D. A., Zedda, M., & Varetto, A. (2018). Physical Practice and Wellness Courses Reduce Distress and Improve Wellbeing in Police Officers. *International journal of environmental research and public health*, *15*(4), 578.

Marks, M. R., Bowers, C., DePesa, N. S., Trachik, B., Deavers, F. E., & James, N. T. (2017). REACT: A paraprofessional training program for first responders—A pilot study. *Bulletin of the Menninger Clinic*, *81*(2), 150-166.

Martin, C. E., Tran, J. K., & Buser, S. J. (2017). Correlates of suicidality in firefighter/EMS personnel. *Journal of affective disorders*, *208*, 177-183.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual review of psychology*, *52*(1), 397-422.

Mathews, K., "Risk management and managerial efficiency in Chinese banks: A network DEA framework"; *Omega*, 41 (2013), 207-215

Mavin, T. J., & Roth, W. M., (2015). Optimizing a workplace learning pattern: a case study from aviation. *Journal of Workplace Learning*, 27:2, 112-127.

McCannel, C. A., Reed, D. C., & Goldman, D. R. (2013). Ophthalmic surgery simulator training improves resident performance of capsulorhexis in the operating room. *Ophthalmology*, *120*(12), 2456-2461.

McClernon, C. K., McCauley, M. E., O'Connor, P. E., & Warm, J. S., (2011). Stress Training Improves Performance During a Stressful Flight. *Human Factors*, 53:3, 207-218.

McCraty, R., & Atkinson, M. (2012). Resilience training program reduces physiological and psychological stress in police officers. *Global advances in health and medicine*, *1*(5), 44-66.

McDevitt, D. M. (2017). *Searching for effective training solutions for firefighting: the analysis of emergency responses and line of duty death reports for low frequency, high risk events* (Doctoral dissertation, Monterey, California: Naval Postgraduate School).

Meehan, M., Insko, B., Whitton, M., & Brooks Jr, F. P. (2002). Physiological measures of presence in stressful virtual environments. *Acm transactions on graphics (tog)*, *21*(3), 645-652.

Moldovanu, R., Târcoveanu, E., Dimofte, G., Lupaşcu, C., & Bradea, C. (2011). Preoperative warm-up using a virtual reality simulator. *JSLS: Journal of the Society of Laparoendoscopic Surgeons*, *15*(4), 533.

Molesworth, B. R.C., & Estival, D. (2015). Miscommunication in general aviation: The influence of external factors on communication errors. *Safety Science*, 73-79.

Moohyun Cha, Soonhung Han, Jaikyung Lee, Byungil Choi (2012), A virtual reality based fire training simulator integrated with fire dynamics data, *Fire Safety Journal*, Volume 50, 2012, Pages 12-24

Nahush A. Mokadam, Richard Lee, Ara A. Vaporciyan, Jennifer D. Walker, Robert J. Cerfolio, Joshua L. Hermsen, Craig J. Baker, Rebecca Mark, Lauren Aloia, Dan H. Enter, Andrea J. Carpenter, Marc R. Moon, Edward D. Verrier, James I. Fann (2015), Gamification in thoracic surgical education: Using competition to fuel performance, *The Journal of Thoracic and Cardiovascular Surgery*, Volume 150, Issue 5, Pages 1052-1058

Navex Global, 2017 Ethics & Compliance Training Benchmark Report.

Nicolaidou, I., Antoniades, A., Constantinou, R., Marangos, C., Kyriacou, E., Bamidis, P., Pattichis, C. S. (2015). A Virtual Emergency Telemedicine Serious Game in Medical

Ong Chao Xiang; Teh Tuan Ann; Chan Ying Hui; Lee Tse Yew (2014), Effectiveness of Gamification in Vocational Technical Education, Proceedings of the European Conference on Games Based Learning, 2014, Vol. 2, pp. 636-644

Orzech, N., Palter, V. N., Reznick, R. K., Aggarwal, R., & Grantcharov, T. P. (2012). A comparison of 2 ex vivo training curricula for advanced laparoscopic skills: a randomized controlled trial. *Annals of surgery*, *255*(5), 833-839.

Ossowski, K. L., Rhee, D. C., Rubinstein, E. N., & Ferguson, B. J. (2008). Efficacy of sinonasal simulator in teaching endoscopic nasal skills. *The Laryngoscope*, *118*(8), 1482-1485.

Oster Jr., C. V., Strong, J. S., & Zorn, C. K. (2013). Analyzing aviation safety: Problems, challenges, opportunities. *Research in Transportation Economics*, 148-164.

Palter, V. N., & Grantcharov, T. P. (2012). Development and validation of a comprehensive curriculum to teach an advanced minimally invasive procedure: a randomized controlled trial. *Annals of surgery*, *256*(1), 25-32.

Palter, V. N., & Grantcharov, T. P. (2014). Individualized deliberate practice on a virtual reality simulator improves technical performance of surgical novices in the operating room: a randomized controlled trial. *Annals of surgery*, *259*(3), 443-448.

Palter, V. N., Grantcharov, T., Harvey, A., & MacRae, H. M. (2011). Ex vivo technical skills training transfers to the operating room and enhances cognitive learning: a randomized controlled trial. *Annals of surgery*, *253*(5), 886-889.

Palter, V. N., Orzech, N., Reznick, R. K., & Grantcharov, T. P. (2013). Validation of a structured training and assessment curriculum for technical skill acquisition in minimally invasive surgery: a randomized controlled trial. *Annals of surgery*, *257*(2), 224-230.

Park, J., MacRae, H., Musselman, L. J., Rossos, P., Hamstra, S. J., Wolman, S., & Reznick, R. K. (2007). Randomized controlled trial of virtual reality simulator training: transfer to live patients. *The American journal of surgery*, *194*(2), 205-211.

Perry, S., Burrow, M. F., Leung, W. K., & Bridges, S. M. (2017). Simulation and curriculum design: a global survey in dental education. *Australian dental journal*, *62*(4), 453-463.

Pia Hoglund (2014), Gamification in Training: engagement motivation, a thesis, *Hanken School of Economics, Helsinki*

Piedmont, R. L. (1998). The Revised NEO Personality Inventory: Clinical and research applications. New York: Plenum Press.

Pilot Training Compass: Back to the future. (2013). *European Cockpit Association*.

Ploch, C. C., Mansi, C. S., Jayamohan, J., & Kuhl, E. (2016). Using 3D printing to create personalized brain models for neurosurgical training and preoperative planning. *World neurosurgery*, *90*, 668-674.

Pokroy, R., Du, E., Alzaga, A., Khodadadeh, S., Steen, D., Bachynski, B., & Edwards, P. (2013). Impact of simulator training on resident cataract surgery. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *251*(3), 777-781.

Potter, B. A., Blickensderfer, E. L., & Boquet, A. J., (2014). Training monitoring skills in helicopter pilots. *Aviation, Space and Environmental Medicine*, 85, 543-549.

Quick, J. A., Vishal Kudav, Jennifer Doty, Megan Crane, Alex D. Bukoski, Bethany J. Bennett, Stephen L. Barnes, "Surgical resident technical skill self-evaluation" *Journal of Surgical Research*, 218 (2017), 144 – 149

Raisinghani, M. S., Chowdhury, M., Colquitt, C., Reyes, P. M., Bonakdar, N., Ray, J., & Robles, J., (2005). Distance Education in the Business Aviation Industry: Issues and Opportunities. *International Journal of Distance Education Technologies*, 3(1), 20-43.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE transactions on systems, man, and cybernetics*, (3), 257-266.

Regambal, M. J., Alden, L. E., Wagner, S. L., Harder, H. G., Koch, W. J., Fung, K., & Parsons, C. (2015). Characteristics of the traumatic stressors experienced by rural first responders. *Journal of anxiety disorders*, *34*, 86-93.

Regehr, C., Glancy, D., Pitts, A., & LeBlanc, V. R. (2014). Interventions to reduce the consequences of stress in physicians: a review and meta-analysis. *The Journal of nervous and mental disease*, *202*(5), 353-359.

Regehr, C., Hill, J., Goldberg, G., & Hughes, J. (2003). Postmortem inquiries and trauma responses in paramedics and firefighters. *Journal of Interpersonal Violence*, *18*(6), 607-622.

Roberts, P. G., Guyver, P., Baldwin, M., Akhtar, K., Alvand, A., Price, A. J., & Rees, J. L. (2017). Validation of the updated ArthroS simulator: face and construct validity of a passive haptic virtual reality simulator with novel performance metrics. *Knee Surgery, Sports Traumatology, Arthroscopy*, *25*(2), 616-625.

Rosa, C., Nadeau, A., Hernandez, E., Kafeero, F., & Zahiga, J., (2016). Delivery strategies to enhance the sustainability of training: lessons from the Food and Agriculture Organization of the

United Nations, International Journal of Training and Development, Vol. 20, Issue: 2, pp 164-176

Rosenbaum, D. P., & Lurigio, A. J. (1994). An inside look at community policing reform: Definitions, organizational changes, and evaluation findings. *Crime & Delinquency*, *40*(3), 299-314.

Roth, W. M., (2015). Flight Examiners' Methods of Ascertaining Pilot Proficiency. *The International Journal of Aviation Psychology*, 25:3-4, 209-226.

Rougeau, K., Chizewski, A., Greenlee, T. A., Greene, D. R., Smith, D. L., & Petruzzello, S. J. (2017). Improvements In Recruit Fitness During A 6-week Firefighter Training Academy: 3709 Board# 156 June 3 9. *Medicine & Science in Sports & Exercise*, *49*(5S), 1064.

Sadasivan, S. and Anand Gramopadhye, "Technology to support inspection training in the general aviation industry", *International Journal of Industrial Ergonomics*, 39 (2009), 608 – 620.

Schauer, P. (Ed.). (2017). *Police Training and Excessive Force*. Greenhaven Publishing LLC.

Schmitz, B., Patrick Felicia, Filippo Bignami, "Acceptance and Use of Game-Based Learning in Vocational Education and Training: an international survey," *International Journal of Game-Based Learning*, 5(4), 50-63, October-December 2015

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, *84*(1), 1.

Schnyder, U. (2014). Treating intrusions, promoting resilience: an overview of therapies for trauma-related psychological disorders. *European journal of psychotraumatology*, *5*(1), 26520.

Schout, B. M., Ananias, H. J., Bemelmans, B. L., d'Ancona, F. C., Muijtjens, A. M., Dolmans, V. E., Scherpbier, A. J. & Hendrikx, A. J. (2010). Transfer of cysto-urethroscopy skills from a virtual-reality simulator to the operating room: a randomized controlled trial. *BJU international*, *106*(2), 226-231.

Schubert, T. W., Friedmann, F., & Regenbrecht, H. T. (1999, April). Decomposing the sense of presence: Factor analytic insights. In *2nd international workshop on presence* (Vol. 1999).

Schwappach, D. L., Blaudszun, A., Conen, D., Ebner, H., Eichler, K., & Hochreutener, M. A. " 'Emerge': benchmarking of clinical performance and patients' experiences with emergency care in Switzerland," *International Journal for Quality in Health Care*, 15(6), 473-485.

Seabrook, R., Brown, G. D., & Solity, J. E. (2005). Distributed and massed practice: From laboratory to classroom. *Applied cognitive psychology*, *19*(1), 107-122.

Sedlack, R. E. (2007). Validation of computer simulation training for esophagogastroduodenoscopy: pilot study. *Journal of gastroenterology and hepatology*, *22*(8), 1214-1219.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, *84*(2), 127.

Shimic, Goran & Jevremovic, Aleksandar & Kostic, Zona & Djordjevic, Dragan. (2015). Assessment based on Serious Gaming Interactive Questions (SGIQ). *Journal of Computer Assisted Learning*, 31, 623–637

Shirai, Y., Yoshida, T., Shiraishi, R., Okamoto, T., Nakamura, H., Harada, T., Nishikawa, J. & Sakaida, I. (2008). Prospective randomized study on the use of a computer-based endoscopic simulator for training in esophagogastroduodenoscopy. *Journal of gastroenterology and hepatology*, *23*(7pt1), 1046-1050.

Silva, G. R., Donat, J. C., Rigoli, M. M., de Oliveira, F. R., & Kristensen, C. H. (2016). A questionnaire for measuring presence in virtual environments: factor analysis of the presence questionnaire and adaptation into Brazilian Portuguese. *Virtual Reality*, *20*(4), 237-242.

Smith, G. M., NewMyer, D. A., Bjerke, E., Niemczyk, M., & Hamilton, R. A. (2010). Pilot Source Study: An Analysis of Pilot Backgrounds and Subsequent Success in US Regional Airline Training Programs. International Journal of Applied Aviation Studies, 10(1).

Sommer, K. J. (2014). Pilot training: What can surgeons learn from it? *Arab Journal of Urology*, 32-36.

Sroka, G., Feldman, L. S., Vassiliou, M. C., Kaneva, P. A., Fayez, R., & Fried, G. M. (2010). Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room—a randomized controlled trial. *The American journal of surgery*, *199*(1), 115-120.

Stanislas, P. (Ed.). (2013). *International perspectives on police education and training*. Routledge.

Stanley, I. H., Hom, M. A., & Joiner, T. E. (2016). A systematic review of suicidal thoughts and behaviors among police officers, firefighters, EMTs, and paramedics. *Clinical psychology review*, *44*, 25-44.

Steinkopf, B., Klinoff, V., Van Hasselt, V., Leduc, T., & Couwels, J. (2016). Behavioral health training for fire rescue personnel. *Fire Engineering*, *169*, 59-67.

Strike Group Defender: The Missile Matrix, *https://www.onr.navy.mil/en/Media-Center/Press-Releases/2015/Strike-Group-Defender-Serious-Game*

SURFnet (2016), The Trend Report 2016 - How technological trends enable customized education, *www.surf.nl/trendreport2016*

Thomas, R. L., (2018). Effect of Active Learning on Instrument Rated Pilots' Knowledge and Self-Efficacy.

Tim Brady, Alan Stolzer (2017), The evolution of air traffic controller training in the United States, Proceedings of International Conference on Information, Communication Technologies in Education, Rhodes, Greece

Torkington, J., Smith, S. G. T., Rees, B. I., & Darzi, A. (2001). Skill transfer from virtual reality

Turner, D. F., (2014). North American Ab Initio Flight Training for Chinese Pilots: A Case Study on Selection. The *International Journal of Aviation Psychology*, 24:2, 83-92.

Training: A Quantitative, Professional Feedback-Informed Evaluation Study. *Journal of Medical Internet Research*, 17(6), e150. *http://doi.org/10.2196/jmir.3667*

Twitchell, S., Elwood Holton, James Trott, "Technical training evaluation practices in the United States", *Performance Improvement Quarterly*, 13(3), (2000), 84-109to a real laparoscopic task. *Surgical endoscopy*, *15*(10), 1076-1079.

Vakharia, V. N., Vakharia, N. N., & Hill, C. S. (2016). Review of 3-dimensional printing on cranial neurosurgery simulation training. *World neurosurgery*, *88*, 188-198.

Van Sickle, K. R., Ritter, E. M., Baghai, M., Goldenberg, A. E., Huang, I. P., Gallagher, A. G., & Smith, C. D. (2008). Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *Journal of the American College of Surgeons*, *207*(4), 560-568.

Vector Solutions, *https://www.vectorsolutions.com/news-media/blog/behind-scenes-vector-solutions-creates-first-kind-virtual-reality-training-firefighters/*

Villard, P. F., Vidal, F. P., Ap Cenydd, L., Holbrey, R., Pisharody, S., Johnson, S., Bulpitt, A., John, N. W., Bello, F. & Gould, D. (2014). Interventional radiology virtual simulator for liver biopsy. *International journal of computer assisted radiology and surgery*, *9*(2), 255-267.

VIRTSIM, *https://www.youtube.com/watch?v=aRWe9xsyReg*

Walker, S. (2005). *The new world of police accountability* Thousand Oaks, CA: SAGE Publications, Inc.

Wampler, R. L., Dyer, J. L., Livingston, S. C., Blankenbeckler, P. L., Centric, J. H., & Dlubac, M. D., (2006). *Training Lessons Learned and Confirmed from Military Training Research* (1850). Arlington, MI: U.S. Army Research Institute for the Behavioral and Social Sciences.

Wentink, M., Stassen, L. P. S., Alwayn, I., Hosman, R. J. A. W., & Stassen, H. G. (2003). Rasmussen's model of human behavior in laparoscopy training. *Surgical endoscopy and other interventional techniques*, *17*(8), 1241-1246.

West, C. P., Dyrbye, L. N., Erwin, P. J., & Shanafelt, T. D. (2016). Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. *The Lancet*, *388*(10057), 2272-2281.

Williams, M. J., Dalgleish, T., Karl, A., & Kuyken, W. (2014). Examining the factor structures of the five facet mindfulness questionnaire and the self-compassion scale. *Psychological assessment*, *26*(2), 407.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225-240.

Wright, B., Dai, M., & Greenbeck, K. (2011). Correlates of police academy success. *Policing: An International Journal of Police Strategies & Management*, *34*(4), 625-637.

Xiang, O. C., Teh Tuan Ann, Chan Ying Hui, Lee Tse Yew, "Effectiveness of Gamification in Vocational Technical Education", *Proceedings of the European Conference on Games Based Learning*, 2, 2014, pp. 636-644

Yi, S. Y., Ryu, K. H., Na, Y. J., Woo, H. S., Ahn, W., Kim, W. S., & Lee, D. Y. (2008). Improvement of colonoscopy skills through simulation-based training. *Studies in health technology and informatics*, *132*, 565-567.

Yiannakopoulou, E., Nikiteas, N., Perrea, D., & Tsigris, C. (2015). Virtual reality simulators and training in laparoscopic surgery. *International Journal of Surgery*, *13*, 60-64.

Yu, F., Hu, X., Ma, C., Zhao, Y., Liu, Y., Yang, F., & Chen, G. (2016, December). MDIS cloth system: virtual reality technology for firefighter training. In *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1* (pp. 219-225). ACM.

Zendejas, B., Cook, D. A., Bingener, J., Huebner, M., Dunn, W. F., Sarr, M. G., & Farley, D. R. (2011). Simulation-based mastery learning improves patient outcomes in laparoscopic inguinal hernia repair: a randomized controlled trial. *Annals of surgery*, *254*(3), 502-509.

Zero Hour: America's Medic, *http://www.virtualheroes.biz/zerohour/*

Z. Xu, X.Z. Lu, H. Guan, C. Chen, A.Z. Ren (2014), A virtual reality based fire training simulator with smoke hazard assessment capacity, *Advances in Engineering Software*, Volume 68, 2014, Pages 1-8.

Appendix A: Terms and Definitions associated with Virtual Reality

Since Virtual Reality has become the major training tool in a wide range of industries, we provide definitions and descriptions of some of the commonly used terms in VR and its application in simulation based learning.

Virtual Environments are environments which are constructed artificially and their design enables them to appear as if these were indeed real.

Immersive Virtual Environments are environments which use gear such as 3-dimensional screens and HMDs such that the environment feels as if it is real to the user.

Fidelity is the extent to which artificial environment created in simulation is similar to actual environment in which the experiment/ training needs to take place. Higher fidelity is usually the product of greater immersion or "real" feeling of the virtual environment.

With greater fidelity, it is expected that learning accuracy is better leading to better training. However, if increasing fidelity is not matched with learning objectives of the program, the results may not be as expected. Moreover, the cost aspect also comes into play. Greater marginal fidelity is usually achieved with greater marginal cost as better fidelity is achieved. The **Alessi hypothesis** states that there is a point beyond which addition of better fidelity does not leads to better training results (Alessi, 1988). Furthermore, the concept of simulation fidelity is not unidimensional, but multidimensional. For example, if the role of simulation is to maximize the transfer of trained behavior to the post training environment, the trainer may select a high fidelity system. However, if the training is carried out to reinforce concepts that have already been acquired and are being regularly practiced, then a low fidelity system may suffice (Beaubien and Baker, 2004). According to Caird (1996), a new training system should have low cost, but still provide trainees an exposure to innovations, develop new skills or learn new procedures maximize the retention and acquisition of knowledge and skills, and yet enable the learners to learn in the most optimal way, allowing them to reflect at ease and incorporate their learning into future effort without causing any hazardous or unsafe conditions.

Virtual environments simulate real life and social interactions through real time graphics on screen, HMDs, body trackers, motion sensors and other such inputs. Immersion creates a sense of presence, which can be measured as through a survey created by Witmer and Singer (1999). Silva et al. (2016) use this questionnaire to describe four main factors involved in measuring presence:

- Involvement: the extent to which one interacts with the environment through focus.
- Adaptation/ immersion: the extent to which the person feels a sense of being enveloped by the surroundings.
- Sensory Fidelity: visual, sensory and haptic perceptions to the senses.
- Interface quality: assessing the control of visual and control interfaces in the environment.

Presence measures

As mentioned previously, Witmer and Singer (1998) published a presence questionnaire, which consists of 7-point rating scale. In addition, there is questionnaire by Schubert et al. (1999) presence questionnaire which distinguishes between spatial-constructive and attention component. Another questionnaire is by Lessiter et al. (2001).

Presence can also be measured by physiological measures. Heart rate, pupil dilation, blink responses and muscle tension are all used to measure presence. However, physiological responses are subjective and different individuals react to same stimuli differently. Barfield and Weghorst (1993), Meehan et al. (2002), Baumgartner et al. (2008), Kober and Neuper (2012) and Kober et al. (2012), Alcaniz et al. (2009) and Clemente et al. (1998) have proposed different techniques to measure physiological responses to presence in immersive virtual environments.

Physical Immersion and devices

When a subject is involved in virtual reality, there needs to be some form of sensory interaction and feedback associated with the virtual environment which can occur if the machine controlling the environment can somehow perceive the presence as seen above and provide feedback. There are several systems used in surgery that provide this kind of experience. We will briefly discuss them here:

Haptics

Systems that interact with the sense of touch. Haptics are divided into two categories – tactile and kinesthetic. Tactile perception consists of texture, vibration and pressure usually through the skin. Kinesthetic perception is through force and movements through muscles and bones in the body. These help develop a deep sense of presence in a user.

Haptic systems are classified as: point based, exoskeletons, wearable systems and locomotive systems.

Point based systems are focused in providing feedback to a user at a single point such as a mouse of a computer or a keypad stroke sound. In surgery Geomagnetic touch, (which allows 6 degrees of freedom), Novint Falcon, (which uses a parallel link mechanism) and Omega 7, (which has a parallel manipulator that connects to a single point); are some examples of point based haptic systems (Figure 7).

Use of point based haptic devices in VR training in surgery

Figure 7 a: Panel A Geomagnetic Touch Haptic Device

Figure 7 b: *Panel B Novint Falcon Haptic Device*

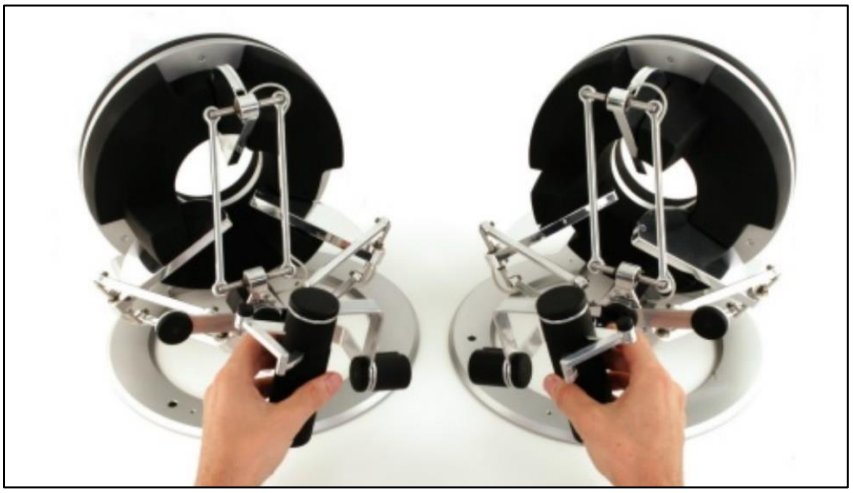


Figure 7 c: *Panel C Omega 7 Haptic Device*

Exoskeleton is a device worn on the exterior and is attached to the subject's body. Exoskeletons provide force feedback to multiple points on the body by stimulating muscles and bones. Examples are CyberGlove (which includes an armature and a hand exoskeleton), X-Arm 2 and ARMin (arm skeleton which provides maximum force and torque to the user) and Novint Xio (sleeve that goes on the arm, and contains a vest, a backpack and a head-mounted display); see Figure 8.

Use of exoskeleton devices in VR training in surgery

Figure 8 a: *Panel A: CyberGlove Exoskeleton*

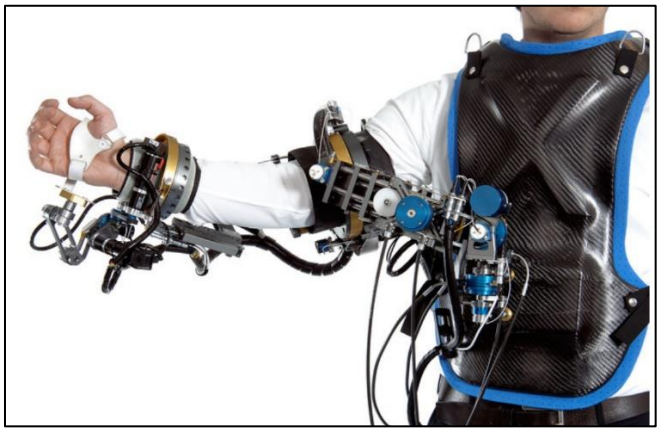


Figure 8 b: *Panel B: X-Arm-2Exoskeleton*

Figure 8 c: *Panel C: ARMin Exoskeleton*

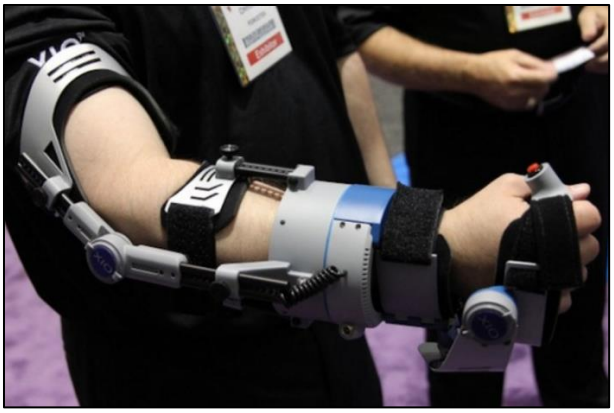


Figure 8 d: *Panel D: Novint Xio*

Locomotive is a full body experience in which the device replicates walking as if the subject were in actual conditions. This has been extensively used in gaming and military. Examples are Tradport (which consists of a visual display, a treadmill and body harness) and Virtuix Omni (Which consists of multidirectional treadmill and can be controlled just by stepping on it); see Figure 9.

Figure 9: Virtuix Omni locomotive haptic system

Motion/ Control is used to create a sense of presence spatially, with involvement and mindfulness of the environment as high fidelity. Since surgical learning is partly focused in motor learning, a lack of control or improper control of motor learning may lead to sub-par training. There are wearable devices, optical trackers, acoustic trackers, mechanical trackers, magnetic and inertial trackers which help in tracking motion/ control of user.

Wearable haptic devices are worn by user such as gloves, exoskeletons etc, as discussed before. For surgery, the cybergrasp tracks the motion of each hand and finger (Figure 10).

Figure 10: *Cybergrasp wearable haptic device*

Optical trackers incorporate a camera system to track motion to create immersive virtual reality. Examples include SELSPOT (Figure 11a), DynaSight sensor (Figure 11b) and Microsoft Kinect (Figure 11c).

Optical Trackers

Figure 11 a: *SELSPOT motion tracker*

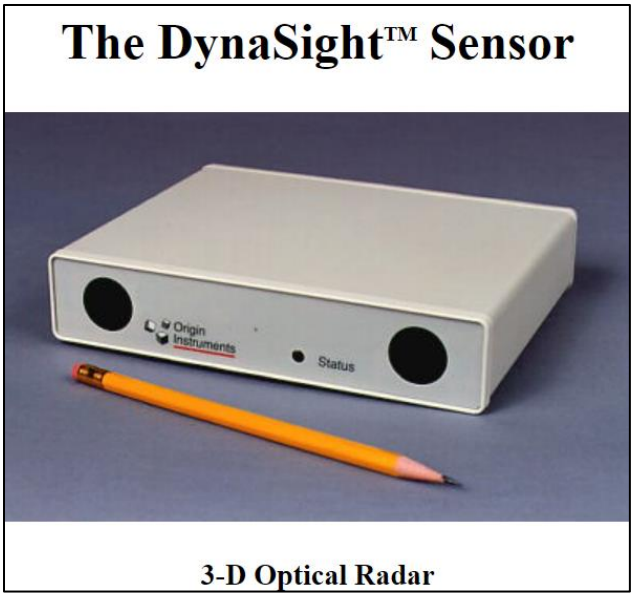


Figure 11 b: *DynaSight Sensor*

Figure 11 c: *Microsoft Kinect*

Acoustic trackers include time of flight and phase coherent trackers. Time of flight tracker has an emitter and receivers which pick up the sound wave produced by the emitter. The phase coherent acoustic trackers determine the difference in phase to compute the distance of the emitter. However disturbances, especially over longer distances can create interference in tracking.

Mechanical trackers use physical connectedness to track the coordinates and track the movement of the object. There are point based mechanical trackers such as the Geomagnetic touch or exoskeleton based trackers such as the X-Arm 2 as discussed previously.

Magnetic trackers use magnetic fields to determine the position and movement of an object. Examples are Polhemus manufactured G4, Patriot etc. (Figure 12a and 12b).

Figure 12 a: *G4 by Polhemus*

Figure 12 b: *Patriot by Polhemus*

Inertial trackers consist of gyroscopes and accelerometers to track motion. The devices have become so common that smart phones track fitness activity with them these days.

Data glove detect joint angles in each finger and therefore transmit this information in the VR system to increase the level of immersion. Example is CyberGlove as seen before.

Eye trackers act as input devices for user's direction of sight and allow the senses to be rendered accordingly. Therefore in surgery, these devices can provide the direction in which the surgery

will proceed and predict a surgeon's response. These are very useful to track a user's response when a complicated situation is occurring. Example of eye trackers are EyeGaze system, BioMuse etc. (Figure 13).

Figure 13: EyeGaze Systems

There are auditory preceptors, which also help increase the immersion. Simple perceptors are headphones and large speakers.

Appendix B: Benchmarking Methodologies, Applications to other industries

The research team at ERAU has conducted extensive study on efficiency analyses and benchmarking of diverse industries. These past successful experiences motivated the team to apply for the FAA's project 'AJI-2 Project #20 Benchmarking Premier Technical Training Providers' under Technical Training and Human Performance (TTHP), Center of Excellence (COE).

It was the intention of the team to utilize their past research experiences and apply quantitative methods to develop benchmarking metrics and key performance indicators (KPI) to technical/professional training in general and aviation technical workforce in particular. However, the scope of the project changed in March 2018 to fully concentrate on non-aviation industries with similar working environment and conditions as air traffic controllers. The new scope was to conduct a literature survey on best practices that these other industries adopt for their professional trainings. The previous sections presented these industries and recommendations on best practices.

The purpose of this section is to provide an overview of the team's past and current benchmarking studies and the methodologies that were adopted for different industries. The team intended to use similar analyses and methodologies to develop performance measurement systems to identify the best practices and benchmark for air traffic controllers and technical training providers. The FAA advised the team that access to data for such an analyses will be challenging and therefore best focus on best practices at other industries.

This section presents some of the team's successful studies and publications on benchmarking and efficiency analyses. These analyses should illustrate the expertise of the team and provide some guidelines on methodologies that we would have implemented to benchmark the air traffic controllers and technical training provided had we have access to data.

This section presents the following:

- An introduction to benchmarking and efficiency analyses.
- Air Transport Research Society - Global Airports Performance Benchmarking Center at ERAU.

- Applications of benchmarking methods to airports and airlines
- Application of benchmarking methods to financial systems.

An introduction to Benchmarking and efficiency analyses.

"What you cannot measure, you cannot manage", Peter F. Drucker.

Benchmarking is a powerful tool to assess a firm or an organization's performance in relation to its goals and objectives, and to evaluate its achievements in relation to peer performance. In the air transport industries, there have been numerous reports and programs benchmarking various aspects of airlines, airports and air navigational service providers. For example, The Technical Operations Performance Improvement and Cost Solutions (TOPICS), developed by Boeing, benchmark airlines' maintenance costs, and identify solutions to maintenance cost issues (Buyers, 2010). Airport Council International (ACI World, 2012) launched A Guide to Airport Performance Measures in 2012 to provide airport managers with an overview and better understanding of the key performance areas of airports, and a set of useful performance measures across a number of categories. ACI-NA has been conducting airport financial benchmarking survey and airport concession benchmarking survey annually, and has also conducted the periodic Airport Compensation & Benefits Benchmarking Surveys. The EUROCONTROL Performance Review Commission (PRC) published its 16th ATM Cost-Effectiveness (ACE) Benchmarking Report in July 2018. These recurrent benchmarking reports by practitioners are often based on ratios that are easy to compute and intuitively easy to understand.

In addition to the recurrent programs, there have been many ad hoc studies benchmarking a specific service provider to its peers. There are also a vast academic literature on airline and airport productivity and efficiency measurement and benchmarking encompassing different methodologies. Yu (2016) provides a survey of alternative methodologies for measuring and comparing productivity and efficiency of airlines. These methodologies have also been applied to airports and air navigational service providers. The following is a brief description of some of the alternative methodologies that are potentially applicable to benchmark air traffic control services.

Partial Factor Productivity and Total Factor Productivity

Partial factor productivity relates a firm or an organization's output to a single input factor. For example, revenue passenger miles per employee is a common airline labor productivity measure, average daily block hours is a common measure of aircraft productivity, and flight hours per air traffic controller hour is a labor productivity measure for air traffic controllers. These partial measures do not require extensive data, and are easy to compute and intuitively easy to understand. One main criticism of partial productivity measures is that productivity of one input depends on the level of other inputs being used. For instance, when an airline or an airport outsources some of its activities, its labor productivity would improve at the expenses of lower productivity of purchased materials and services. Another problem with partial productivity measures in the air transport industries is that airlines, airports, and air navigational service providers generate multiple outputs, thus a single output measure, such as revenue passenger miles, do not reflect an airline's total output. Nevertheless, partial productivity measures can provide useful insights and practical guidance for improving performance when firms or organizations operate under similar environments, or over time within the firm or organization.

Total Factor Productivity (TFP) is defined as the amount of aggregate output produced by a unit of aggregate input. TFP became a popular approach to measure airline performance in the early 1980s. The main issue with TFP is how to aggregate the outputs and inputs. A commonly accepted way to aggregate outputs and inputs is the translog multilateral index procedure developed by Caves, Christensen and Diewert (CCD, 1982). This CCD Index procedure allows researchers to compute a consistent output (input) index that provides a consistent comparison across firms or organizations and over time within a firm or organization. The typical inputs (or outputs) aggregation formula can be specified as follows:

$$Ln\, Z_k - Ln\, Z_j = \sum_i \frac{R_{ik} + \overline{R_i}}{2} ln \frac{Z_{ik}}{\widetilde{Z_i}} - \sum_i \frac{R_{ij} + \overline{R_i}}{2} ln \frac{Z_{ij}}{\widetilde{Z_i}} \qquad (1)$$

where Z_k is the aggregate index of input (or output) for kth observation, Z_{ik} is ith input (output) for kth observation, the R_{ik} are weights, a bar over a variable indicates the arithmetic mean and a tilde over a variable indicates the geometric mean. Revenue shares are often used as the weights in

output aggregation (with the assumption of constant returns to scale), while cost shares are used as weights in input aggregation. Each observation represents a specific firm or organization at a particular time. This procedure allows transitive comparisons across all observations (across firms and over time within a firm) via a series of binary comparisons between each observation and the means of the data.

TFP comparison for observation j and observation k will then be:

$$Ln\,TFP_k - Ln\,TFP_j = (\ln Y_k - \ln Y_j) - (\ln X_k - \ln X_j) \qquad (2)$$

Where, Y represents outputs; and X represents inputs.

TFP is essentially the weighted average of partial factor productivities of all inputs a firm or organization uses. Since TFP requires detailed data on revenues and expenses that are not always readily available to researchers, they have lost some popularity in recent years. Other methodologies, such as data envelopment analysis (DEA), that do not require detailed financial data, have since become more popular.

Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a non-parametric method for measuring productive efficiency of "decision making units" (DMUs) with multiple inputs and multiple outputs. In their seminal paper on DEA, Charnes, Cooper and Rhoades (1978) developed a mathematical programming approach that combines and transforms multiple inputs and outputs into a single efficiency index. DEA has been used as a benchmarking tool in the air transport industries, including airlines and airports. It does not require detailed data on revenues and expenses, and does not assume a particular functional form for production frontiers. The efficiency index is computed through a set of linear programs, and is essentially the ratio of the total weighted output to the total weighted input. The weights are determined by the linear programming optimization. DEA allows each DMU to select the weights that maximize its own efficiency score. Generally, DMUs will place higher weights on the inputs that they use the least and on the outputs they produce the most. In this sense DEA shows each DMU in its best possible light. A score of 1

indicates that the DMU is efficient relative to other DMUs in the sample, and a value less than l indicates a DMU being inefficient.

Econometric Methods

The Econometric methods require the specification of production or cost functions with constant parameters. The ***Conventional Econometric Methods*** involve the estimation of a production or cost function. The estimated production or cost function is then used to identify changes in productivity or efficiency. The production function generally includes a time trend variable, and the estimated coefficients of the time trend are used to measure the rate of technical progress that in turn indicates the rate of productivity growth.

Since airlines and airports produce more than one output, and there tends to be high collinearity between the inputs, many studies estimate cost functions instead of production functions. Similar to the production functions, a time variable is generally included in the cost functions to estimate the growth of output and technical progress. One problem with estimating cost functions is that it requires data on input prices that are not always readily available.

Conventional econometric estimation of production and cost functions implicitly assumes that all firms are on the efficient frontier. ***Frontier Production or Cost functions***, on the other hand, recognize that some firms may not be on the efficient frontier, thus a firm's efficiency is measured based on its location relative to the frontier. Within the frontier function frame, a distinction is made between two different methods: the deterministic frontier method and the stochastic frontier method. The main difference between the two methods is that the deterministic method attributes all deviations from the frontier to inefficiency while the stochastic method distinguishes the deviations into a random component capturing statistical noise and an inefficiency component.

The ***Deterministic Frontier Methods*** assume that all firms share a common production frontier, and all variations in firm performance are attributed to variations in firms' efficiencies relative to this common frontier. This assumption ignores the very real possibility that a firm's performance may be affected by factors entirely outside its control, such as bad weather, labor disputes, and so on, as well as by factors under its control (inefficiency). ***Stochastic Frontier Models*** were developed to distinguish the effects of "noise", including exogenous random shocks, measurement

errors, misspecification of production functions, and etc, from inefficiency. The basic idea behind the stochastic frontier method is that the deviation from the frontier is composed of two parts. A symmetric component permits random variation of the frontier across firms and captures the effects of measurement error, statistical "noise", and exogenous random shocks outside the firm's control. A one-sided component captures the effects of inefficiency. The stochastic frontier methods have been widely applied over the last three decades using various estimation techniques

Air Transport Research Society - Global Airports Performance Benchmarking Center at ERAU

The Air Transport Research Society (ATRS) was established in 1995 to enhance the research capability for multi-national and multi-disciplinary issues on air transportation, and to foster interaction between international and national aviation academics dealing with policy, management strategy and infrastructure issues.

The ATRS Annual Global Airport Performance Benchmarking Project was initiated in 2000. The project is now hosted at College of Business at Embry-Riddle Aeronautical University (ERAU) in Daytona Beach, FL (USA). Dr. Yu has been a member of the ATRS Airport Benchmarking Task force since its inception, and is the research director for the benchmarking project.

The ***ATRS Global Airport Benchmarking Project*** measures and compares the performance of several important aspects of airport operations: Productivity and efficiency, unit costs and cost competitiveness, financial results and airport charges. The report also examines the relationships between various performance measures and airport characteristics as well as management strategies in order to provide a better understanding of observed differences in airport performance. The 2018 report includes 205 airports and 24 airport groups of various sizes and ownership forms in Asia Pacific, Europe and North America. The following is an excerpt from Part I of the 2018 ATRS report:

Productivity and Efficiency: The report first measures and compares a number of partial measures of productivity, including labour productivity, capital productivity, and the so-called "soft cost" input productivity. Soft costs input is a catch-all input other than labour and capital costs, including costs of outsourced services, consultant services, utility costs, travel expenses, non-labour building and equipment maintenance, and other purchased materials and services. The soft cost inputs account for 29.2% to 93.2% of airports' non-capital expenditure, depending on the airport. Therefore, it is extremely important to consider this soft cost input in examining airport performance.

In this report, the partial productivity measures of all inputs including soft cost inputs are computed and compared using different output measures, and examined in relation to various airport characteristics. It should be noted that the partial productivity measure is influenced by the levels of other inputs being mixed in the production process. For example, an airport's labour productivity depends on how much of its services (such as ground handling, fire and police, security screening, snow removal, etc.) are outsourced to other firms or suppliers because the number of employees they need would depend on the extent of outsourcing. Consequently, labour productivity or other partial factor productivities alone are not good indicators for comparing efficiency of airport operations among airports. There is a need to construct an aggregate measure of productivity for all inputs airports use. Variable Factor Productivity (VFP) is used in this report as airport's overall productivity indicator in the short to medium term. VFP index is computed by aggregating labour productivity and soft cost input productivity using variable cost shares as the weights for aggregation, and it measures how efficiently an airport utilizes variable inputs for a given level of capital infrastructure and facilities. The differences in observed VFP across airports are decomposed into various sources including differences in airport business environments (airport size, passenger-freight traffic mix, percentage of international passengers, average aircraft size), and a measure of capacity shortage for runways and/or terminals (which is a proxy measure for an airport's service level to air carriers). After removing the effects of the factors beyond airport managers' control, we compute and compare the "true" managerial efficiency of the airports in the form of "residual" VFP.

Unit Costs and Cost Competitiveness: The report presents a number of unit cost measures, focusing on variable cost as there is no established measure for capital costs that is consistent and comparable across different countries. These include the commonly used unit cost measures, such as labour cost per passenger, labour cost per aircraft movement, etc. In addition, we developed an aggregate output index for airport services by aggregating all types of outputs and services including passenger traffic volume, aircraft landings and takeoff movements, and non-aeronautical services such as concessions, rentals and development activities, etc. This output index represents a single aggregate measure of all outputs each airport produces and services. A unit cost index is then constructed as cost

per unit of this aggregate output. These unit cost measures are then examined in relation to various airport characteristics variables and airport service levels.

Observed unit cost differences do not reflect true comparative cost competitiveness between airports, as they operate under different operating and regulatory environments. What determine cost competitiveness are input prices paid by an airport and how efficiently the airport operates. Therefore, a cost competitiveness indicator is constructed by summing the effects of input price and efficiency.

Financial Results: The report presents a number of revenue generation indicators and commonly used financial ratios; including aeronautical revenue per aircraft movement, total revenue per passenger, return on equity, current ratio, return on asset, etc.

Airport Charges: Airports derive revenues through a combination of aeronautical and non-aeronautical (commercial) activities. Aeronautical charges are levied on airlines for the use of an airport's runway, apron and terminal facilities, whereas non-aeronautical revenues are derived from a growing range of commercial activities, some of which are performed by the airports themselves and others are sub-contracted. It is difficult to form a common platform on which to compare the non-aeronautical charges across the airports. Therefore, this report focuses on aeronautical charges.

Applications of benchmarking methods to airports and airlines

The following provides examples of the research team's past work on airport or airline benchmarking and efficiency analyses.

Size versus Efficiency - A case study on U.S. Commercial Airports

Bazargan, M., Vasigh, S. (2003), "Size versus Efficiency - A case study on U.S. Commercial Airports", *Journal of Air Transport Management,* Vol. 9, pp 187-193.

This study analyzed 45 US commercial airports (the top fifteen large, medium, and small hub airports) to identify benchmarks and key performance measures for their efficiencies. In defining the airport system in terms of efficiency, two categories of performance measures are discussed: the financial and the operational measures. The Federal Aviation Administration (FAA) classifies airports based on the percentage of national total of passengers enplaned. According to FAA, a large hub airport is defined as the one with 1% or more of national enplaned passengers. These percentages for medium and small hub airports are 0.25 to 0.99% and 0.05 to 0.24% respectively. Based on this classification, we attempted to address the following questions:

- Is there any evidence that efficiencies of these three hubs are different?
- If so, what hubs tend to be more efficient?
- Are these efficiencies stable over time or they tend to fluctuate annually?
- What (if any) are the factors that make the hubs inefficient? What improvements are necessary to make these hubs more efficient?

To answer these questions we identified top 15 airports from each hub (total 45 airports). We collected the input and output data over 5 years. The airports were selected in a way that consistently retained their positions in the three hubs grouping in all the five-year study period.

We adopted the following arrays of input and output data for the airports over the 5 years of study. It should be noted that these input and output data are based on the total annual number of operations including air carriers, general aviation, air taxis and military.

Four input measures were selected for this study as follows:

- Operating expenses – represent the financial resources needed to run an airport. These figures include: personnel compensation and benefits, communications and utilities, supplies, materials, repairs and maintenance, services and other expenses.
- Non-operating expenses - includes debt services, capital expenditures and other non-operating expenses.
- Number of runways - The available number of runways at each airport for the year of study.
- Number of gates - Includes all the gates with jet ways and other non-jet-way gates.

Six output data were collected for each airport over the five-year study period as follows:

- Numbers of passengers - These figures represent passengers arriving or departing at a specific airport.
- Number of air carrier operations - Represents the total number of air carrier movements (landings and take offs). We adopt the same interpretation of the air carriers as referenced by FAA.
- Number of other operations - Includes all the movements other than air carriers such as commuters, general aviation and military.
- Aeronautical revenue - Includes all revenues that are generated by aviation activities such as landing fees, terminal fees, apron charges, fuel flowage, fixed base operators (FBOs), rentals and utilities.
- Non-Aeronautical revenue - Includes rent, concessions, parking, rental cars, catering and so on.
- Percentage of on time operations - Represents the percentage of total number of operations (both air carriers and others operations) on time divided by total number of operations at the specific airport annually.

Data Involvement Analyses (DEA) methodology was adopted for this study. The analyses identified that:

- There exists a significant difference among the efficiencies of all three hubs
- Small hubs consistently outperform medium and large hubs in their relative efficiencies

- Recommendations for each airport on how to improve on their inefficiencies.

Incorporating Negative Externalities Into Productivity Assessments of US Airports

- **Transportation Research A**, Volume 62, April 2014, Pages 39–53, by Davide Scotti, Martin Dresner, Gianmaria Martini, and Chunyan Yu

The study analyses the efficiency of 44 US airports for the period 2005-2009. To be technically efficient, an airport must maximize its outputs given the infrastructure, facilities and other resources (inputs). Airports are generally considered to "produce" three physical outputs in the form of traffic flows: passengers, cargo, and aircraft movements. However, as traffic increases, the likelihood of flight delays also increases. Flight delays have negative impacts on passengers and shippers' satisfaction, that is. Furthermore, airport activities generate noise and air pollutions as byproducts that are not desirable by passengers, shippers or the general public. This study incorporates three negative or undesirable outputs in assessing airport efficiency.

The study uses a directional distance function approach to estimate the efficiency scores, then conducts a second stage analysis to examine the impacts of a set of factors, including fleet mix, airport size, share of night flights, and share of international passengers, etc, on airport efficiency. The purpose of the second stage analysis is two-folder: (1) To test the effect of each variable on determining airport efficiency scores; and (2) to check if the determinants exert a different effect on efficiency after the introduction of undesirable outputs into the production process.

The results show that estimated efficiency scores are sensitive to the inclusion of the undesirable outputs (externalities). The implications are that the inclusion of these externalities into the calculation of efficiency may impact managerial decisions. The results further suggest the existence of economy scale in airport operations.

Ownership Forms Matter for Airport Efficiency: a Stochastic Frontier Investigation of Worldwide Airports,

- **Journal of Urban Economics,** Vol. 64, Issue 2, September, 2008, 422-435, by Jia Yan, Tae H. Oum, and Chunyan Yu

The study develops an advanced form of stochastic frontier model to examine how different ownership forms affect airport efficiency based on an unbalanced panel of 109 airports around the world. The sample airports represent different size, ownership and institutional arrangements.

A translog stochastic cost frontier function was estimated via a Bayesian approach, and it was jointly estimated with variable input share equations to improve the efficiency of estimation. The cost frontier function includes three outputs (passengers, aircraft movements, and non-aeronautical services), two fixed inputs (runways and terminal size), and two variable inputs (number of employees and non-labor-variable input). The non-labor-variable input includes all expenses not directly related to capital or labor inputs. To examine the effects of ownership forms on airports, the study classified airports into seven categories including majority private ownership, government –private mixed ownership with government majority, independent airport authority, government corporations, directly operated by government departments, shared ownership by multiple governments, and quasi-public port authorities. In addition, a number of airport characteristics variables were considered in the model as well, including percentage of international passengers, percentage of cargo traffic, etc.

The key findings are (a) airports owned and/or controlled by majority private firms, autonomous public corporations or independent authorities are more efficient than those owned and/or controlled by government branch (city/state), multiple level governments, or US ports authorities; (b) there is an almost 100% probability that airports controlled/ operated by independent airport authorities are more efficient than those controlled/operated by US port authorities, and there is 93% probability that US city/state run airports are more efficient than those operated by US port authorities; (c) there is about 80% probability that airports owned/operated by a majority private firm achieve higher efficiency than those owned/operated by the mixed enterprise with government majority ownership; and (d) airports owned/operated by a Government controlled agencies (US ports authorities, shared government ownership, US city or state

government, mixed enterprises with government majority ownership) have significantly lower efficiency in multiple airport markets than in single airport markets.

New Evidences on Airline Efficiency and Yields: A Comparative Analysis of Major North American Air Carriers and its Implications

- **Transport Policy**: Volume 12, Issue 2, March 2005, 153-164, by Tae H Oum, Xiaowen Fu, and Chunyan Yu

This paper measures and compares the productive efficiency, unit cost competitiveness, and average yields of 10 major full service carriers in Canada and the United States for the period of 1990–2001. The study considers five output variables and five input variables. The output variables are scheduled passenger services (in Revenue Tonne Kilometers – RTK), scheduled freight services (in RTK), mail service (in RTK), non-scheduled services (RTK), and incidental services output. The inputs are labour, fuel, materials, flight equipment, and ground property and equipment (GPE).

The outputs (inputs) are aggregated to form a multilateral output (input) index using the translog multilateral index procedure. Subsequently, the total factor productivity (TFP) of the sample airlines were computed as the ratio of the aggregate output over the aggregate input. The TFP is considered as a 'gross" productivity measure as it is likely influenced by factors beyond the control of airlines, thus it may not reflect the true productive efficiency of the airlines. A set of log-linear TFP regressions were estimated on a number of airline network characteristics variables, including average stage length, composition of outputs, average load factor, rate of fleet capacity adjustment, output size, and year dummies. The regression analysis has two objectives: (1) to identify the potential effects of variables on gross TFP, and (2) to compute a residual TFP index after removing the effects of the variables beyond managerial control. The residual TFP is considered as an indicator of airlines' productive efficiency. The study further constructed an unit cost competitiveness index by subtracting input price index from the residual TFP index.

The results indicate North American airlines achieved substantial improvement in productive efficiency during the 1990-2001 period, especially in the mid-1990s, which allowed them to cope

with rising input prices and competitive pressure on yields. The results further show that airlines' financial success depends not only on its cost efficiency, but also heavily on its pricing and yield management strategy. The comparison between airlines in the United States and airlines in Canada find that although Canadian carriers were considerably less efficient than their US counterparts, they were able to maintain comparable cost competitiveness due to their lower input prices resulting from the devaluation of the Canadian dollars.

Application of benchmarking methods to financial systems

Misvaluation and Behavioral Bias in Financial Markets

This study analyzes pockets of inefficiencies in stocks trading in the U.S. financial markets between the period 1992 and 2010. The paper actually builds upon the stochastic frontier model laid down by Aigner, Lovell and Schmidt (1977) and Meeusen and van Den Broeck (1977). The models propose the conventional regression equation to estimate the best fit to a given sample of data. However the model introduces an additional term, called the one-sided error. Applied to the financial world, particularly to any security 'i' at time 't', the mathematical expression can be represented as:

$$R_{it} = \beta_0 + \beta_1 R_{mt} + u_{it} + \mu_{it}$$

$$\mu_{it} \sim N^{+}\left(0, {\sigma_\mu}^2\right) \; or \; N^{-}\left(0, {\sigma_\mu}^2\right)$$

$$\varepsilon_{it} = u_{it} + \mu_{it}$$

Here, R_{it} is the return from stock

R_{mt} is the return from market (denoted as 'm')

u_{it} is the two-sided error term which follows a random walk (and under certain regular conditions, follows a normal distribution)

μ_{it} is the one-sided error with half normal distribution with expected value of zero and variance ${\sigma_\mu}^2$.

ε_{it} is the composite error term accounting for imperfections and inefficiencies in the markets

This one-sided error term can be modelled as half-normal, exponential or truncated normal distributions. This enables a researcher to form an envelope around the data points depending on the side the frontier is chosen. If production frontier is chosen, then the model estimates a surface of maxima for the dependent variable, implying the best the production unit can do, given inputs. If, on the other hand, a cost function is chosen, then the model estimates a surface of minima, implying that is the best the unit can do to minimise cost. What this one-sided error therefore, does is to the model the inefficiency depending on the objective chosen. For a production function, the nature of inefficiency will automatically make $\mu_{it} \geq 0$. For a cost function, $\mu_{it} \leq 0$. We use this phenomenon in stock markets in our 2015 paper (Gokhale et al., 2015).

Our study employs the premise that if there is no behavioral bias in the markets (i.e. $\varepsilon_{it} = u_{it}$) and the overall model has an error term that follows a random walk. However, if there is overvaluation bias, then the distribution for the one-sided error term does not have an expected value of zero. Instead, if there are consistently better than expected returns from stocks as compared to the predicted values from the market model posited by the efficient market hypothesis (EMH), then we refer to this situation as overvaluation bias. In such a scenario, we use a stochastic frontier model that has positive bias in the expected value of the stock returns. With overvaluation, μ_{it} is distributed as nonnegative half normal, $N^+(0, \sigma_\mu^2)$. Similarly, If there is undervaluation, μ_{it} would be distributed as nonpositive half normal, $N^-(0, \sigma_\mu^2)$. Actual returns fall short of expected fundamental returns by $|\mu_{it}|$.

$$E(R_{it}) = R_{it} - \mu_{it} \text{ and } R_{it} = E(R_{it}) + \mu_{it},$$

where $\mu_{it} < 0$. Notice that when $\sigma_\mu \to 0$, the distribution collapses to a spike at zero so that $\mu_{it} \to 0$. There is no undervaluation in this case and the market model is the correct specification. Thus, the null hypothesis of interest is $\sigma_\mu = 0$, which we test using a one-sided likelihood ratio test (Coelli, 1995). If the null hypothesis is rejected, then μ_{it} is negative (i.e., ε_{it} is skewed in the negative direction), implying undervaluation.

In the 2015 paper, we deduce the index of over/ undervaluation and develop a method to compute the index for Toyota Motor Company. For several years in the early 2000s the company had been growing at a furious pace. Along with that, its stock price grew by 120% in the five years from 2002 – 2007, a period in which it surpassed General Motors as the largest car manufacturing company. We hypothesized that Toyota's stock prices may have been found with overvaluation bias during this period. This hypothesis is based on the premise that furiously growing companies grab more attention of investors, who tend to be greatly optimistic about its future performance. Since stock price is a function of the discounted future cash flows, it is plausible for investors to make this hypothesis based on their forecasted cash flows of Toyota based on its furious growth. However, this confidence of investors came tumbling down because of the recall of millions of vehicles due to sticky accelerator pedal. This led to a decline in market value of the company and as Gokhale et al. (2014a) find, the company lost over 19% of its market value in terms of cumulative abnormal returns. We deduce the expected value of σ_μ for Toyota before and after the event and determine that the company indeed was overvalued in pre-event period. On the other hand, after the recalls were made, the company was neither overvalued nor undervalued. We define a measure of this inefficiency obtained from frontier based analysis as the fundamental valuation index or FVI as follows $FVI = \frac{\sum_{t=1}^{T} \widehat{\mu_{it}}}{T}$. We obtained the estimate of FVI for Dow 30 companies and invested in the companies with negative and significant FVI for undervaluation. We keep rebalancing this portfolio every year. We find that the Dow 30 companies which are strongly undervalued tend to produce better than expected risk adjusted rates of return, as compared to the major stock indices (the Dow Jones Industrial Average and the Standard and Poor's 500 Average).

We extend this study to the companies of the Standard and Poor's 100 companies from the year 1990 through the year 2010 (see Gokhale et al., 2014b), identifying the companies which are undervalued (FVI for undervaluation which is significant at either 1% or 5% level), investing in them for a period of 1 year and then selling them in favor of buying stocks which were undervalued that year at our threshold significance levels. Using this strategy, we identify undervalued stocks for the period of study and find that stocks which are undervalued produce better risk adjusted returns than both the DJIA and S&P 500 Indices with returns of 405% and 575% respectively for 1% and 5% undervaluation bias as compared to 280% returns from S&P 500 and 315% from the DJIA.

References

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, *6*(1), 21-37.

Coelli, Tim, “Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis,” *Journal of Productivity Analysis*, 6 (3), 1995, 247-268.

Gokhale, J., Tremblay, C. H., & Tremblay, V. J. (2015). Misvaluation and behavioral bias in financial markets. *Journal of Behavioral Finance*, *16*(4), 344-356.

Gokhale, J., Brooks, R. M., & Tremblay, V. J. (2014a). The effect on stockholder wealth of product recalls and government action: The case of Toyota's accelerator pedal recall. *The Quarterly Review of Economics and Finance*, *54*(4), 521-528.

Gokhale, J., Schroeder, E., & Tremblay, V. J. (2014b). Valuation Bias and Profit Opportunities in Financial Markets. *International Journal of Economics and Finance*, *6*(4), 147.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International economic review*, 435-444.

Other Publications in the field

Buyers, Tom (2010), "Optimizing Airplane Maintenance Economics", Aero Quarterly, QTR_01.10. pp 5-9

https://www.boeing.com/commercial/aeromagazine/articles/qtr_01_10/pdfs/AERO_Q1-10_article02.pdf

ACI World (2012), "ACI Launches a Guide to Airport Performance Measures", ACI Media Release, February 27, 2012

https://aci.aero/news/2012/02/27/aci-launches-a-guide-to-airport-performance-measures/

Chunyan Yu (2016), "Airline Productivity and Efficiency: Concept, Measurement, and Applications", forthcoming in **Advances in Airline Economics** – Volume 5, edited by John Bitzan, James Peoples and Wesley Wilson, Emerald Group Publishing Limited, Bingley, UK