Title: Part Task and Variable Priority Training in First Year Anesthesia Resident Education: A combined didactic and simulation based approach to improve management of adverse airway and respiratory events.

Author Listing: Ken B Johnson, M.D., Noah Syroid, M.S., Frank Drews, Ph.D., Lazarre Ogden, M.D., David Strayer, Ph.D., M.D., Diane Tyler, R.N., Julia White, R.N., and Dwayne Westenskow, Ph.D.

Departments and Institution: Departments of Anesthesiology. Cognitive Psychology, and Biomedical Engineering, University of Utah, Salt Lake City, Utah, United States.

Corresponding Author:
Ken B. Johnson, M.D.
Associate Professor
Learned Family Associate Professorship
Department of Anesthesiology
University of Utah
30 North, 1900 East, Room 3C444
Salt Lake City, Utah 84132-2304
Phone: (801) 205-1039
FAX: (801) 581-4367
Email: ken.b.johnson@hsc.utah.edu

Reprints: Request for reprints should be with the corresponding author.
Summary: The feasibility of using part task and variable priority training to improve the response of first year anesthesia residents in their management of seven adverse airway and respiratory events was explored. Clinical performance in managing these adverse events was assessed prior to and after the 12 month training period. Part task and variable priority training improved first year resident management of these adverse events.

Acknowledgements: Funded in part by a research grant from the Anesthesia Patient Safety Foundation, Indianapolis, IN

Short Title: Part Task and Variable Priority Training for Adverse Airway and Respiratory Events
ABSTRACT

Introduction: Part task training (PTT) focuses on dividing complex tasks into small components followed by intensive concentrated training on each individual component. The goal of PTT is to reduce processing demands when performing multiple complex tasks simultaneously by developing levels of automaticity. Variable priority training (VPT) focuses on optimal distribution of attention when performing multiple tasks simultaneously. The goal of VPT is to allow resource allocation to be priority driven. The aim of this study was to explore how principles of PTT and VPT adapted to clinical training in anesthesia implemented over a twelve month period would improve first year anesthesiology residents’ management of simulated adverse airway and respiratory events when compared to conventional training methods. We hypothesized that participants with PTT and VPT would perform better than those with standard training.

Methods: After IRB approval, 22 first year University of Utah anesthesia residents were enrolled. In addition to clinical training in the operating room, participants were randomly divided into two groups for additional training. Participants assigned to the control and experimental groups received standard didactic training and VPT and PTT of equal duration over 12 months respectively. Participant ability to manage seven different adverse airway and respiratory events was assessed prior to and after the 12 month training period. Five metrics of resident performance were measured. These included number of correct tasks, percent making a correct diagnosis, time to a correct diagnosis, assessment of cognitive workload, and an assessment of scenario comprehension.

Results: After the 12 month training period, participants in the experimental group were able to complete more appropriate monitoring and treatment tasks and answered more comprehension questions correctly than those in the control group. There was no
difference in perceived workload, the time to a correct diagnosis, or the number of
correct diagnoses between groups.

**Conclusion**: This study confirmed in part our study hypotheses. Our results suggest that
VPT and PTT are promising adjuncts to didactic and simulation based training for
management of adverse airway and respiratory events.
INTRODUCTION:

Part task training (PTT) and Variable Priority Training (VPT) are tools that have been developed by psychologists to optimize human performance while simultaneously asked to complete multiple tasks where some of the tasks are time sensitive. These techniques have been successfully implemented in a number of simulator based professional training arenas and have led to higher pass rates in settings where students are asked to manage multiple tasks.\cite{Gopher,1994}

PTT is defined as the break down of large multi-component tasks into a set of component tasks that then by training of individual components either separately or in various combinations can become highly automatized \cite{Whightman,1985,Lintern,1991,Damos,1991}. A consequence of this training is a reduction of the processing demands of large complex multi-component tasks by streamlining effort associated with the individual elements of the task. In addition, focused training on individual components lead to more rapid development of automatic skills that might otherwise not be achieved in the context of the whole task.

VPT is a method for training people to flexibly distribute attention over multiple aspects of a task. Participants in VPT learn to coordinate and control how attention is allocated to components of a task and assign different processing priorities to the components as they are performed in concert. For example, Gopher \cite{Gopher,1982,Gopher,1989} used VPT to instruct novice pilots how to monitor instruments more efficiently. Under dynamic flying conditions, pilots learned to vary their priorities on how they attended to the information presented on an aircraft console. VPT required pilots to distribute their attention across several sources of information and avoid focusing on any single source of information yielding a more efficient detection of problems and a lower drop-out rate than with novice pilots receiving standard training. Similarly, Brikner &
Gopher {Brickner, 1981 #15} found that VPT reduced performance decrements when compared to standard training when task difficulty varied. VPT appears to foster a flexible cognitive style that reduces the likelihood of cognitive tunnel vision.

The Committee on Professional Liability of the American Society of Anesthesiologists (ASA) has maintained a database of anesthesia malpractice claims since 1985. Closed claim analysis of adverse events in anesthesia has revealed that airway and respiratory events are the largest class of injuries (37%) identified in the database.{Morray, 1993 #7; Cheney, 1991 #8; Caplan, 1990 #6} Analyses of this database demonstrated that the adverse airway and respiratory events were made up of primarily unrecognized esophageal intubations and a result of difficult intubations{Caplan, 1990 #6}. Additional categories related to adverse airway and respiratory events included airway trauma, pneumothorax, airway obstruction, aspiration, and bronchospasm{Cheney, 1991 #8}.

More recent work analyzing the closed claim database has pointed out that adverse airway and respiratory events occur with a higher frequency in pediatric patients with more severe consequences (e.g. higher rate of mortality or brain injury){Morray, 1993 #7}. In these analyses, anesthetic care was judged as less than appropriate in 44% of the adult claims and 54% of the pediatric claims. Furthermore, these complications were felt to be preventable with better monitoring in approximately one half of the claims. Airway management difficulties, impaired vigilance, inadequate supervision, poor judgment, diversion of attention, and misinterpretation and misuse of data were also noted as potential sources for bad outcomes associated with adverse airway and respiratory events.{Morray, 1993 #7}

Given that adverse airway and respiratory events often occur during a time period where anesthetists perform multiple tasks at once, clinicians can quickly get
overwhelmed when attempting to formulate an accurate diagnosis and appropriate treatment plan in a timely fashion. Skills acquired through PTT and VPT may improve vigilance in detecting and responding to these types of life threatening events.

We employed PTT and VPT techniques as part of the didactic and simulation based training for first year anesthesia residents (CA1’s) over a 12 month period. Training was directed towards optimizing detection, diagnosis, and appropriate treatment for some of the more frequently encountered adverse airway and respiratory events reported in the closed anesthesia malpractice claims database. The aim of this study was to explore how PTT and VPT would improve first year anesthesiology residents’ management of simulated adverse airway and respiratory events when compared to conventional training methods. We hypothesized that residents in the experimental group (with PTT and VPT) when compared to those in the control group would (1) complete more critical tasks essential to treating an adverse event, (2) reach a correct diagnosis more often and more quickly, (3) report a decreased perception of workload and (4) demonstrate an increased level of comprehension when managing simulated adverse airway and respiratory events.
METHODS:

After Internal Review Board Approval, 22 University of Utah anesthesia residents were consented to participate in a between-subjects twelve month study. Enrollment was confined to residents in their first year of clinical anesthesia training. First year training consisted of 12 month long rotations in general adult anesthesia, pediatric anesthesia, and surgical intensive care practice in a university setting. Participants were randomly divided into two equal size groups: the control and experimental groups. All training activities, to include operating room supervision, simulation training, and didactic sessions were conducted by board certified or board eligible staff anesthesiologists.

Using data from a prior simulator based study using a similar experimental design {Johnson #17}, a power analysis was performed to determine the number of subjects required to detect difference in the time to diagnosis or the number of correct treatment actions between the control and experimental groups. A statistically significant sample size of 10 was calculated based on a type I margin of error set to 0.05 and 80% power to detect a difference between groups{Machin, 1987 #16}.

Baseline Evaluation

To establish a baseline skill level in managing adverse airway and respiratory events, participants were asked to care for seven simulated patients using an adult and a pediatric human simulator (HPS version 5.55, METI, Sarasota, FL.). The seven adverse airway and respiratory topics used in these simulations are presented in Table 1. In addition to the patient simulator, a physiologic monitor (Datex AS/3, Helsinki, Finland) displayed (i) the electrocardiogram, pulse oximeter, and capnogram waveforms and (ii) digital values for heart rate, blood pressure, oxygen saturation, end-tidal carbon dioxide, and fraction of inspired oxygen. The pulse oximeter tone was provided. All
standard monitor alarms were set to default (factory) limits. Mechanical ventilation was provided by an anesthesia machine (Narcomed 2B, North American Dräger, Telford, Pennsylvania). A peripheral nerve stimulator was attached to the mannequin to assess the extent of neuromuscular blockade.

A scenario vignette and a simulated patient record and pre-anesthetic evaluation were prepared for each scenario. In a quiet room adjacent to the simulation laboratory, participants reviewed the scenario vignette and pre-anesthesia evaluation. Participants were then introduced into the simulation laboratory. Participants were encouraged to talk aloud during the evolution of each scenario. Video and audio information were recorded for all simulations. A video image of the physiologic monitor was inset into a video image of the participant caring for the simulated patient. Also superimposed of the video image was a timer with a resolution to 1 second. The order of the seven scenarios was randomized for all participants. Each scenario was programmed to last seven minutes from the start of the adverse event.

During each scenario, scenario report forms were completed by two investigators watching the video images in a room adjacent to the simulation laboratory. Investigators were blinded to participant group assignment. Scenario report forms were developed by two board certified anesthesiologists (co-authors of this manuscript) and contained diagnostic and therapeutic tasks appropriate for each adverse event (Table 1). A series of three pilot studies were conducted with volunteer residents not included in either study group. Through these pilot studies, the scenario report forms were refined to contain: (1) the start of each adverse event, (2) the time to identifying the appropriate diagnosis, (3) a checklist of 110 diagnostic and therapeutic tasks, and (3) the time an appropriate diagnosis was made.
Completed scenario report forms were compared between investigators. For time discrepancies of greater than 1 minute or data entry omissions (i.e. one investigator indicated a diagnostic or therapeutic event occurred but the other did not), a third blinded investigator reviewed the video image to reconcile the difference.

Following each scenario, participants completed a questionnaire regarding their comprehension of the adverse event. An example of the comprehension questions for scenario #3 are presented in (Table 2). Participant also completed a self assessment of their perceived physical and cognitive workload using a National Aeronautic Space Administration Task Load Index (NASA-TLX) questionnaire{Hart, 1988 #3}. The NASA-TLX evaluated six areas: mental demand, physical demand, temporal demand, self assessment of performance, self assessment of effort and self assessment of frustration.

As previously described Syroid et al, Anesthesiology 2002 Mar;96(3):565-75, a composite score was estimated following each scenario.

For each scenario, five metrics of performance were measured: (1) the number of diagnostic and/or treatment tasks completed to properly manage the simulated respiratory events, (2) the percentage of correct diagnoses, (3) the time required to reach a correct diagnosis, (4) an assessment of perceived workload, and (5) an assessment of comprehension of the adverse simulated airway or respiratory event.

**Between Groups Interventions**

Training interventions were of identical duration for both study groups. Four areas of clinical competence in managing an adverse airway and respiratory events were identified by the investigators as: (1) the ASA Difficult Airway Algorithm{, 2003 #19}, (2) a differential diagnosis for hypoxia, (3) treatment options for each item in the differential diagnosis, and (4) a readily available knowledge of normal ranges of
cardiopulmonary variables and key relationships between selected variables. Study
sheets were created for each of these areas and were distributed to all participants at
the beginning of the 12 month intervention period. [Consider an appendix – I don’t think
it is necessary].

*Didactic Training*

All participants received a series of didactic sessions in a classroom setting and
simulation sessions equally spaced out over the 12 month intervention period.
Participants assigned to the control group received five 90 minute simulation sessions
and forty 45 minute didactic sessions (grand rounds, case conferences, textbook chapter
reviews {Morgan, 2002 #20}, and visiting professor lectures). The chapter reviews
included topics covering cardiopulmonary physiology, airway management for both adult
and pediatric patients, recognition of and management for a difficult airway, and
administration of anesthesia to patients with cardiopulmonary disease.

Participants in the experimental group received fifteen 45 minute PTT sessions
and twenty five didactic sessions as described for the control group. The PTT sessions
were conducted in a classroom setting with a personal computer for each participant.
The content of the PTT sessions included reviewing and committing to memory items on
four study sheets described above.

As a study aid, computerized flashcards were developed to review information on
the study sheets using computer programming software (Java, Sun Microsystems Inc.,
Santa Clara, CA). A series of questions were created for each sheet and entered into a
question pool. A personal computer was used to automate question presentation and to
record response time and accuracy. The flashcards were designed to (1) be time
limited, (2) randomly re-introduce questions that were answered incorrectly into the
question pool, (3) present the participant with the correct response when an incorrect answer was entered, (4) record the number of questions answered correctly, and (5) present the participant with the number of correctly answered questions at the end of the timed flashcard study session. The intent of the flashcard sessions was to improve participant ability to recall information from the study sheets in a time sensitive manner.

[Consider placing a link to a URL that contains the flashcard tests – Do-able Noah?]

Three types of flashcards were developed. The first type, “Fill in the Blank” was designed to introduce information from the four study sheets by having participants look up answers to questions and then fill in blanks. During the first three sessions, participants were allowed to consult the study sheets when entering their answers. Subsequently, participants were asked to fill in the blank from memory. The second type, “Qualitative Assessment” was designed to have participants provide a qualitative evaluation from memory of values presented in the study sheets. For example, qualitative responses of high, low, or normal were obtained when presented with normal and abnormal values for selected vital signs or responses of correct or incorrect for questions exploring statements regarding pulmonary pathophysiology and management of a difficult airway. The intent of this flashcard type was to (1) accelerate clinical interpretation of physiologic monitor and ventilator values encountered in the perioperative environment and (2) recall of information relevant to an adverse airway or respiratory event. The third type, “Patient Management” was designed to put into practice utilization of information contained within the four study sheets in management of simulated adverse respiratory or a difficult airway events. In this type of flashcards, participants were presented with sets of cardiopulmonary variables and were asked to comment on their state (high, low, or normal) and identify the most likely diagnosis and treatment consistent with the variable profile.
During each session, 25 minutes were allocated to use of computerized flashcards. The remaining time was dedicated to small group discussions where participants were asked to recall information from the four study sheets from memory in the presence of their peers and discussion proctor. The overall goal of this PTT was to achieve a level of mastery of the material on the study sheets such that material contained within them would be easily recalled during stressful moments of clinical care as can occur during management of critical airway or respiratory event.

**Simulation Training**

All participants received five 90 minute simulations over the 12 month intervention period. The simulation sessions covered topics on the difficult airway, management of hypoxia, hypertension, tachycardia, bradycardia, and hypovolemia. All scenarios were pre programmed into the simulator computer. Simulated patient information (patient description, pre operative evaluations, operating room records, laboratory studies) and teaching objectives were standardized for each simulation training session to ensure that all participants were exposed to the same adverse events in each scenario. Participants were divided into smaller subgroups of 5 to 6 participants to facilitate simulation based training. In each simulation training session, 5 to 6 brief scenarios (10 minutes) were conducted. This allowed each participant to manage at least one adverse cardiopulmonary event per 90 minute training session while the remaining participants observed. Following each scenario, a short debriefing (5 minutes) was conducted to review the participant’s performance.

In the control group, instruction during these five simulation sessions was targeted on reviewing the teaching objectives for each simulation topic. Participants in the experimental group received VPT adapted to management of adverse airway and
respiratory events. VPT consisted of asking participants to review four areas of patient
data: (1) pertinent findings from the patient history, (2) a brief targeted physical exam
(airway, breathing, and circulation), (3) physiologic data, and (4) mechanical ventilation
data (if available). A list of items for each area is presented in Table 3.

The goals of VPT adapted to adverse airway and respiratory events were: (1) to
ensure that participants considered all available data in a timely manner (i.e. in less than
two minutes) and (2) to provide a means of more in depth consideration of
cardiopulmonary function should there be an abnormal finding. This was accomplished
by considering: (1) the relationship between airway pressures (peak and plateau) and
the set and delivered tidal volume to identify potential abnormalities in airway resistance
and lung compliance, and (2) the relationship between the inspired oxygen content
(FiO₂) and the arterial PaO₂ and oxygen saturation (SpO₂) to identify potential perfusion
ventilation mismatches.

During each simulation scenario, the participant caring for the simulated patient
was asked to consider, from memory, the check list and key relationships described in
Table 3 while the remaining participants had a paper copy of the checklist in front of
them and checked off items mentioned by the participant managing the adverse event.
During the debriefing, the participant’s performance in evaluating the items listed on
Table 3 was reviewed. Thus, during each simulation session, each participant critiqued
4 to 5 other participant’s ability to evaluate the items in Table 3 and had one opportunity
to use the check list from memory while managing an adverse event. Starting with the
third simulation session, in addition to recalling pertinent patient data, participants were
asked to generate a differential diagnosis and perform additional diagnostic and
therapeutic maneuvers.
Post Training Performance Assessment

Following the twelve month training period, each participant underwent a simulation based assessment of their skill in managing seven adverse airway and respiratory events as described for the baseline assessment. The vignettes used to introduce the participant to each scenario were changed but the adverse events remained the same as the ones used in the baseline analysis.

Statistical Analysis

Data from each of the 7 scenarios were compiled into five data sets, one for each of the five performance metrics. A repeated measures analysis of variation was used to compare groups over time for each metric. P values less than 0.05 were considered significant. Data are presented as mean ± standard error of the mean.
RESULTS

Of the 22 participants enrolled, one participant did not complete the study. Midway through the 12 month intervention period, the participant elected to pursue post graduate training in a different specialty other than anesthesiology. 11 (3 female 8 male) and 10 (3 female 7 male) participants were randomly assigned to the control and experimental groups respectively. The remaining participants all completed the baseline assessment, the twelve month intervention (five 90 minute simulation based training sessions and the forty 45 minute didactic sessions), and the post intervention assessment.

Baseline and post intervention assessments of the five performance metrics (number of properly completed diagnostic and treatment tasks, the percentage of correct diagnoses, the time to a correct diagnosis, the composite NASA-TLX scores, and the percentage of correct comprehension questions) are presented for all participants by group in Figures 1-4. At baseline, there was no significant difference between groups in any of the five performance metrics. Following 12 months of training, participants in both groups demonstrated a significant change in all five performance metrics when compared to baseline (P value < 0.0001).

Participants in the experimental group completed more appropriate diagnostic and therapeutic tasks than participants in the control group (P = 0.0210). As illustrated in Figure 1, out of 110 possible tasks, participants in both groups completed on average 34 appropriate tasks at baseline while managing the 7 adverse event scenarios. After the 12 month training period, both groups showed an increase in the number of tasks completed, but participants in the experimental group completed on average 10 more appropriate tasks than their control group counterparts.
As presented in Figure 2, participants in the experimental group tended to provide a correct diagnosis more often than the control group (a mean difference of 11% after training). Similarly, participants in the experimental group tended to require less time to identify the correct diagnosis than the control group (a mean difference of 0.5 minutes after training). The repeated measures of analysis of variance, however, revealed no difference between groups for either of these metrics (P value = 0.448 and 0.359 respectively).

Participants reported via the NASA TLX a decrease in their self assessment of perceived workload between baseline and post training (Figure 3). There was no difference between the experimental and control groups (P value = 0.2589). Following the 12 month training period, participants in the experimental group completed more appropriate diagnostic and treatment tasks when managing the adverse airway.

Percentages of correct responses to a comprehension survey following each adverse airway or respiratory event are presented in Figure 4. The repeated measures analysis of variance reported a significant difference between groups (P < 0.0001). Participants in both groups completed approximately 50% of the questions correctly at baseline but, following the 12 month training period, the experimental group completed more questions correctly than participants in the control group.
DISCUSSION

In this study we set out to employ two novel approaches (PTT and VPT) to resident education directed at managing adverse airway and respiratory events. Our hypotheses were that PTT and VPT when used over a 12 month period would improve selected metrics of clinician performance when managing life threatening adverse airway and respiratory events. Our results in part confirmed our hypotheses. Participants in the experimental group (with PTT and VPT) out performed participants in the control group in two of five performance metrics and in the remaining three, there was no difference between groups.

The scenarios used to assess their skills in managing a difficult airway or adverse respiratory were designed to quickly develop life threatening conditions as realistically as possible. This was done to mimic the rapid decline in physiologic function that often accompanies an adverse event of this nature and to examine participant skill while working under stressful conditions. Each scenario was designed to run its course within 7 minutes of a triggering adverse event. During several scenarios, where participants had difficulty identifying the correct diagnosis or therapy, seven minutes was perceived by participants as a long stressful period. With this in mind, we never allowed the simulated patient to expire even though the current management would have likely led to a poor outcome. Our approach to assessing participant ability to effectively manage adverse events was to use five metrics of performance. Three of these metrics were measured by investigators monitoring participant performance in an adjacent room. The other two were in the form of questionnaires completed following each adverse event scenario.

Overall, participants following 12 months of training improved in their management of the seven simulated adverse events. Participants completed more
critical tasks, made the diagnosis more often and more quickly, reported a decrease in their perceived workload, and reported a higher level of comprehension with regard to the adverse events they managed.

**Critical Task Completion**

With regard to the first metric, completion of appropriate diagnostic and therapeutic tasks, participants in the experimental group following a year of training with PTT and VPT were able to complete on average 10 more of tasks than those in the control group. To put these results in perspective, two board certified anesthesiologists, after reviewing the content of the seven adverse events identified a total of 110 tasks that they deemed appropriate in managing these events. Not all of the items on the check list were mandatory for a successful outcome. The task lists were designed to encompass several approaches to managing these events when appropriate. Some of the tasks were critical to effectively manage the adverse events (i.e. recognize an esophageal intubation) and others were important, but one of several would suffice (i.e. administering a beta agonist versus a potent inhaled agent in managing a bronchospasm following emergence from anesthesia). Prior to training, participants were able to complete 31% of these tasks. After a year of training, participants completed 40 and 50% of these tasks in the control and experimental groups respectively. The clinical implication of this increase is that participants in the experimental group had the wherewithal to maintain their focus and systematically assess and treat an adverse event in a stressful environment. We suggest that this is a function of their training that focused on an efficient yet complete survey of all available information, and using easily retrieved mental templates to organize and use this information to implement therapeutic interventions.
Diagnostic Accuracy and Time to Diagnosis

The PTT and VPT implemented over the 12 month intervention period did not result in any detectable difference between groups in both the percentage of participants that made a correct diagnosis and the time required to make a correct diagnosis. For both metrics, subtle trends suggested that PTT/VPT may improve diagnostic performance but these trends were not significant. With regard to the percentage of correct diagnoses (Figure 2a), the improvement was large from baseline (35-39% correct) to the end of 12 months of training (61-73%) independent of group assignment making detection of a difference between groups challenging. It is also important to point out that in several instances; correct diagnoses were mentioned amidst participants verbalizing a list of potential diagnoses. It was difficult to ascertain whether participants were guessing or were confident of their diagnosis. The time to diagnosis (Figure 2b) also demonstrated an improvement from a mean time of just over 5 minutes to 3 to 3.5 minutes. When interpreting these results, it is important to point out that scenarios were terminated if participants were unable to make a diagnosis within 7 minutes. Had participants been allowed to continue trying to identify the correct diagnosis indefinitely, although beyond the time frame of what a patient would be expected to survive, the means would have been shifted to higher values indicating that the training in either group was not as good as the data would suggest.

Perceived Workload

The NASA TLX was used to assess the perceived workload following each scenario. We found that the perceived workload following 12 months of training decreased by approximately 30% from baseline. This is an expected finding given that
participants, with a year of clinical experience are more likely to have a better subjective belief about their performance than after one week of residency. PTT and VPT made no impact on the perceived workload. One nuance of this result is that additional work generated by participants using cognitive aids developed through PTT and VPT did not contribute to increasing the perceived workload.

Scenario Comprehension

Scenario specific questionnaires were used to identify participant understanding of the adverse events occurring during each scenario. We found that at baseline, participants had a limited understanding of the adverse events answering on average just above 50% of the questions correctly. Following 12 months of training, participants in the control group had a modest increase in their mean percentage of questions answered correctly (60%) whereas the experimental group had a large increase in questions answered correction (73%). These results illustrate how PTT and VPT allowed participants to collect and retain more pertinent information in regard to the adverse event. These results may be a direct result of our findings in Figure 1 where participants in the experimental group completed more diagnostic and therapeutic tasks. In so doing, they gained a better understanding of the adverse events and were able to answer more questions correctly.

Training Time

Each participant during the 12 month intervention period typically spent 65 to 70 hours a week for 49 weeks (3100 to 3400 hours) in a supervised clinical practice setting. By contrast, each participant spent 76.5 hours in didactic training over the 12 month intervention period, of which 37.5 hours were directed at PTT/VPT in management of a
difficult airway or adverse respiratory event. These hours in didactic training only represent 2.5 and 1.2% respectively of their overall training time during the intervention period. It is important to point out that although participants may have observed or participated in the care of a handful of life threatening adverse events during their first year of training, the amount of time spent in managing these events is most likely very small (i.e. perhaps an additional 1% of their overall training time). Despite a relatively short period of training time, the impact of mismanagement during these events is harmful and costly. Although implementation of this study led to more time spent on educational effort directed at managing adverse events, spreading ten PTT and five VPT events over twelve months led to long gaps in time between training sessions. With the goal of improving recall automaticity, shortening the time between training sessions may have improved training efficiency. We recognize that ensuring adequate non clinical time to complete this type of education effort can be difficult. Nevertheless, it is the opinion of the authors that more time (i.e. more than 3-4%) should be allocated to this type of training to achieve an appropriate level of skill in managing these types of events.

**Study Limitations**

There are several limitations to this line of investigations. As suggested by Gaba (Gaba, 2002 #14), there are several limitations to this line of investigation primarily related to the use of simulation based evaluations to assess clinician performance during an adverse event. Difficulties arise in designing scenarios that mimic real world complex patients with all their inherent unpredictability yet minimize the impact of inter-participant variability, rater variability, and over simplification of test measures making study results less likely to reflect the impact of novel teaching tools on clinical performance. In addition, participants often act in ways they wouldn't normally behave during a real
adverse event. For example, participants may be hyper-vigilant in anticipation of an “impending crisis” or exhibit a cavalier attitude where participants elect to behave in a more relaxed fashion than they would in during a real adverse event [Wong AK Can J Anesth 2004; 51(5): 455-464.]. For the most part, participants in our study typically fell into the hyper-vigilant mode as they worked their way through the seven scenarios.

Another limitation of this investigation is that we used simulation based evaluation tools in place of more conventional means of assessing clinical performance. Conventional methods used to measure resident clinical performance typically include monthly written evaluations by faculty members, performance on in-service and mock oral exams, attendance at didactic training sessions, and individualized discussion at semiannual clinical competency meetings. We felt that using conventional assessment tools are too subjective, difficult to measure, and would most likely not be sensitive enough to detect improved skills in managing adverse airway or respiratory events learned through PTT and VPT.

Since simulation based evaluation tools are not routinely used, the potential benefit of PTT and VPT outside of our experimental protocol may go undetected. This makes educational activities that utilize PTT and VPT or other similar simulation based techniques and their intense educational overhead less attractive to implement until more widespread use of higher resolution assessment tools become more commonplace and used in determination of clinical competence.

An additional limitation of this study was that although we found significant improvements in implementing PTT and VPT techniques, the long term benefit of these training techniques beyond the year long training period is unknown.

Another limitation is that in order to ensure that all participants were treated in an identical manner, each participant received the same scenarios at the beginning and end
of the year. The scenario vignettes were changed, but the nature of the adverse events were identical. We observed no overt evidence of recollection of the baseline testing after one year (i.e. no participant stated that (s)he had seen that scenario before). If there was such an influence, we anticipate that it would be present in both groups which potentially limits the generalizability of the results to clinical practice, but has no effect on evaluating the efficiency of PTT and VPT.

In summary, the intent of this study was to explore how the use of PTT and VPT over a 12 month period would influence first year anesthesia residents’ ability to manage adverse airway and respiratory events. The intent of this training was to develop a flexible cognitive style that can increase vigilance and produce a high level of automaticity during critical events. We found PTT and VPT techniques adapted to the management of adverse airway and respiratory events to be effective training tools for first year anesthesia residents.
**Table 1.** Scenario titles and acceptable diagnostic and therapeutic tasks for each scenario.

**Adult Scenario #1: Unanticipated Difficult Airway**

1. Detects enlarged tongue
2. Detects swollen posterior pharynx
3. Recognize low SpO₂
4. Recognizes inadequate ventilation
5. Recognizes an unanticipated difficult airway
6. Calls for help
7. Ensure FiO₂ set to 100%
8. Attempt laryngoscopy (blade enters mouth)
9. Mask ventilation following 1st laryngoscopy attempt
10. 2nd Laryngoscopy attempt (different blade enters mouth)
11. Mask ventilation following 2nd laryngoscopy attempt
12. Place LMA and inflate cuff
13. Awaken patient

**Adult Scenario #2A: Bronchospasm on emergence**

14. Auscultates pulmonary fields
15. Detects decreased breath sounds with wheezing
16. Recognizes low SpO₂
17. Detects high peak airway pressure
18. Detects low tidal volume
19. Recognizes severe bronchospasm
20. Calls for help
21. Ensure FiO₂ set to 100%
22. Suctions oral airway
23. Attempts manual ventilation
24. Administers intravenous epinephrine
25. Considers administration of subcutaneous terbutaline, lidocaine and/or ketamine
26. Administers a sedative hypnotic and muscle relaxant
27. Performs laryngoscopy and endotracheal intubation and inflates cuff
28. Administrers beta agonists inhaler
29. Uses inhaler circuit adapter
30. Administers potent inhaled agent

**Adult Scenario #2B: Development of a tension pneumothorax**

31. Auscultates pulmonary fields
32. Detects decreased breaths sounds on the left hand side
33. Recognizes low SpO₂
34. Detects high peak airway pressure
35. Detects low tidal volume
36. Detects presence of a left tension pneumothorax
37. Calls for help
38. Ensure FiO₂ set to 100%
39. Performs a needle decompression of left chest
40. Provides mechanical ventilation

**Adult Scenario #3: Aspiration following induction**

41. Auscultates pulmonary fields
42. Detects course breath sounds with crackles
43. Recognizes low SpO₂
44. Once intubated, detects high peak airway pressures
45. Once intubated detects low tidal volume
46. Calls for help
47. Ensure FiO₂ set to 100%
48. Places patient in trendelenberg and turns head
49. Suctions oral airway
50. Attemps mask ventilation
51. Administers a sedative hypnotic and muscle relaxant
52. Laryngoscopy and intubation to include inflation of the endotracheal tube cuff.
53. Suctions endotracheal tube
54. Provides mechanical ventilation
55. Set positive end expiratory pressure to 5 cm H₂O
56. Consider intravenous furosemide 5 to 20 mg

**Pediatric Scenario #1: Esophageal intubation**

57. Checks endotracheal tube depth
58. Checks endotracheal tube cuff pressure
59. Laryngoscopy to validate endotracheal tube placement
60. Auscultates pulmonary fields
61. Detects absence of breath sounds
62. Recognizes low SpO₂
63. Recognizes minimal or no end tidal CO₂
64. Detects no tidal volume with manual ventilation
65. Recognizes esophageal intubation
66. Calls for help
67. Ensure FiO₂ set to 100%
68. Remove esophageal tube
69. Mask ventilate to improve oxygen saturation
70. Perform laryngoscopy and intubate trachea to include inflating cuff
71. Suction stomach

**Pediatric Scenario #2: Laryngospasm, bronchospasm, No intravenous access**

72. Auscultates pulmonary fields
73. Detects decreased breath sounds and wheezing
74. Recognizes low SpO₂
75. Recognize tachycardia
76. Recognizes minimal or no end tidal CO₂
77. Detects high peak airway pressures
78. Detects no tidal volume with manual ventilation
79. Recognizes inadequate ventilation
80. Recognizes laryngospasm
81. Recognizes bronchospasm
82. Calls for help
83. Ensure FiO₂ set to 100%
84. Attempts manual ventilation
85. Administers subcutaneous epinephrine or terbutaline
86. Administers intramuscular succinylcholine and atropine
87. Performs laryngoscopy following succinylcholine
88. Intubates trachea and inflates cuff
89. Once intubated, administers beta agonist inhaler
90. Uses circuit adapter for inhaler
91. Suctions endotracheal tube

Pediatric Scenario #3: Faulty ventilator circuit

92. Checks endotracheal tube depth
93. Checks endotracheal tube cuff pressure
94. Laryngoscopy to validate endotracheal tube placement
95. Auscultates pulmonary fields
96. Detects normal breath sounds
97. Recognizes low SpO₂
98. Recognizes minimal or no end tidal CO₂
99. Detects low peak airway pressure (< 10 cm H₂O)
100. Detects low tidal volume (< 500 mL)
101. Recognizes inadequate ventilation
102. Recognizes anesthesia circuit has an air leak
103. Replaces ventilator circuit
104. Uses Jackson Reese Circuit to ventilate patient
105. Calls for help
106. Ensure FiO₂ set to 100%
107. Increases tidal volume
108. Increases respiratory rate
109. Increases inspiratory flow rate
110. Increases oxygen fresh gas flow
Table 2 Comprehension questions for each scenario

**Adult Scenario #1: Unanticipated Difficult Airway**

1. The room air oxygen hemoglobin saturation was (Low, Normal, High).
2. The arterial pO2 was (Low, Normal, High).
3. The relationship between the room air oxygen saturation and arterial pO2 is (Normal, Suggestive of an intrapulmonary shunt, Suggestive of an enlarged Function Residual Capacity).
4. Structures or conditions that impaired successful intubation of this patient include all of the following: (Tongue, Posterior pharynx, Laryngospasm, Bronchospasm, Nasopharyngeal Airway).
5. Transtracheal jet ventilation would have been a logical next step in airway management (True, False).

**Adult #2: Bronchospasm on emergence and development of a tension pneumothorax**

1. Following extubation, airway resistance was (Low, Normal, High).
2. Following re-intubation, the patient developed (Endobronchial intubation, Fat embolism, Negative Pressure Pulmonary Edema, An increase in the vital capacity, None of these).
3. Following re-intubation, the peak airway pressure was (Low, Normal, High) and the measured tidal volume was (Low, Normal, High).
4. Prior to extubation, the depth of the endotracheal tube was (Deep, Normal, Shallow).
5. Auscultation of lung sounds after re-intubation revealed (Course breath sounds, wheezing, absent breath sounds, normal breath sounds).

**Adult #3: Aspiration following induction**

1. Once intubated, the delivered tidal volume was (Low, Normal, High).
2. The breath sounds were (Normal, Coarse, Wheezing, Absent).
3. The pulmonary compliance was (Low, Normal, High).
4. The airway resistance was (Low, Normal, High).
5. Following aspiration, broncho-alveolar lavage is warranted (True, False).

**Pediatric Scenario #1: Esophageal intubation**

1. The peak airway pressure following intubation by the anesthesia resident you were supervising was (Low, Normal, High).
2. The endotracheal tube cuff was intact (True, False).
3. The end tidal CO2 following intubation by the anesthesia resident you were supervising was (Low, Normal, High).
4. The delivered tidal volume following intubation by the anesthesia resident you were supervising was (Low, Normal, High).
5. The most likely source of hypoxia was (Inadequate ventilation, bronchospasm, laryngospasm, pulmonary embolism, carboxyhemoglobin).

Pediatric Scenario #2: Laryngospasm, bronchospasm, No intravenous access

1. Following intubation, the peak airway pressure was (Low, Normal, High).
2. Bronchospasm is an example of (Dead Space, Intrapulmonary Shunt, Chest Wall Rigidity, Aspiration of foreign body).
3. Following intubation, the delivered tidal volume was (Low, Normal, High).
4. The inspiratory flow rate was set at (Low, Medium, High).
5. The patient became hypoxic because of a ventilator malfunction (True, False).

Pediatric Scenario #3: Faulty ventilator circuit

1. The peak airway pressure was (Low, Normal, High).
2. The delivered tidal volume was (Low, Normal, High).
3. The inspiratory flow rate was set at (Low, Medium, High).
4. The endotracheal tube cuff was intact (True, False)
5. The wall oxygen source was low (True, False).

Correct answers are presented in Bold.
Table 3. Variable priority training checklist

**History (check if considered)**
Findings related to intraoperative patient course
Findings related to potential therapeutic interventions

**Physical (check if considered)**

**Airway**
- Suction ETT
- Secretions (Present / Absent)
- Check ETT cuff pressure (High / Normal / Low)
- Check ETT depth (Deep / Normal / Shallow)
- Check for ETT disconnect (Connected / Disconnected)

**Breathing**
- Evaluate chest excursion (Equal Bilaterally / Unequal / None)
- Auscultate lung sounds (Normal / Crackles / Wheezes / None)

**Circulation**
- Check pulses (Present / Absent)
- Evaluate skin color (Normal / Abnormal)
- Estimated blood loss (Normal / Abnormal)

**Physiologic data (check if considered)**
- Heart Rate H / N / L
- Heart Rhythm (NSR / ST / AFib / PVC / VT / VFib / Asystole)
- SpO$_2$ H / N / L
- Blood Pressure H / N / L
- End Tidal CO$_2$ H / N / L
- Respiratory Rate H / N / L
- Temperature H / N / L
- PA pressures H / N / L
- Wedge pressure H / N / L
- CVP H / N / L

**Mechanical data (check if considered)**

**Settings**
- FiO$_2$
- Respiratory Rate
- Tidal Volume

**Measured**
- FiO$_2$ H / N / L
- Respiratory rate H / N / L
- Tidal volume H / N / L
- Peak Inspiratory Pressure H / N / L

**Key Relationships**
- Relationship of peak airway pressure to tidal volume:
  - High Compliance / Normal / Low Compliance
- Relationship of FiO$_2$ to oxygen saturation: Saturations
  - Higher /Normal/ Lower than expected with given FiO$_2$

**Use of VPT:** Excellent / Good / Fair / Poor

ETT indicates endotracheal tube, H / N / L indicates High / Normal / Low, NSR, ST, AFib, PVC, VT, and VFib indicate normal sinus rhythm, sinus tachycardia, atrial fibrillation, premature ventricular contraction, ventricular tachycardia, and ventricular fibrillations respectively, FiO$_2$ indicates inspired oxygen content.
FIGURE LEGENDS

**Figure 1.** Mean number of appropriate diagnostic and treatment tasks completed while managing seven simulated adverse airway or respiratory events (110 tasks possible within the 7 scenarios). Data are presented by group at baseline and following the 12 month training period.

**Figure 2.** Panel A: Mean percentage of correct diagnoses made while managing simulated adverse airway and respiratory events by group at baseline and following the 12 month training period. Panel B: Mean time to reaching a correct diagnosis by group at baseline and following the 12 month training period. The maximum allowed time for each scenario was 7 minutes.

**Figure 3.** Mean of reported composite NASA TLX scores of perceived workload by group at baseline and following the 12 month training period.

**Figure 4.** Mean percentage of correct responses to a questionnaire reviewing participant's comprehension of each scenario’s adverse event. Data are presented by group at baseline and following the 12 month training period.
Figure 1

Baseline

Number of Completed Tasks

Control Group

Experimental Group

Post Training

Number of Completed Tasks

0
10
20
30
40
50
60
Baseline
Post Training

DRAFT
Figure 2

Panel A

Baseline
Post Training

Percent Correct

Control Group
Experimental Group

Panel B

Baseline
Post Training

Time to Diagnosis (min)
Figure 3

![NASA TLX Score Graph]

- **Baseline**
  - Control Group
  - Experimental Group

- **Post Training**
  - Control Group
  - Experimental Group

- **Y-axis**: NASA TLX Score
- **X-axis**: Baseline, Post Training
Figure 4

Baseline Percent Correct

Control Group
Experimental Group

Post Training

Percent Correct

0 20 40 60 80

Baseline Post Training