UNIVERSITY OF NORTH DAKOTA

A study evaluating if targeted training for startle effect can improve pilot reactions in handling unexpected situations.

DR. MICHAEL GILLEN, PH.D.

Disclaimer

I would like to thank the FSF, United Airlines, and specifically Captain Michael Quiello for their support of my research. The opinions expressed in this presentation are formulated from my doctorial research and dissertation. These opinions are mine alone and may not reflect those of United Airlines or the FSF.

Goals in Research

- Identify an industry problem
- Use live and active pilots to explore solutions
  - Generally airline pilots
    - Survey
    - Simulator performance
  - Historically difficult group to do research on
- Develop/explore and validate possible mitigations
  - Perspective as a researcher, operator , and instructor.

Topic

- Emergencies in aircraft often involve high-stress decision-making, which must be accomplished correctly in real (often limited) time, with incomplete or confusing information.
- Crews are often startled at the onset of such events.
  - Unfortunately, incorrect initial decisions at the start of an emergency often result in delayed aircraft recovery and in some cases lead to an undesired aircraft state (UAS).
  - Decisions in stressful environments are often made with information from past experiences, training, and pattern matching (Rasmussen, 1983).
- Although each emergency is surrounded by unique circumstances, training over a broad array of scenarios and circumstances may give flight crews enough background information to manage the situation for a successful outcome.

Problem Statement

- Recent airline accidents point to a rapid degradation from controlled flight following an unusual event when the flight crew becomes startled.
- There has been very little training among airline crews on how to successfully manage a sudden and often stressful event that requires quick and accurate decision-making (BAE, 2012).
- Accident data has indicated that when an incorrect decision is made, the likelihood of a successful outcome decreases (Hilscher, Breiter, & Kochan, 2012).
- This study sought to determine if specific and targeted training could help mitigate the effects of flight crews being startled by implementing a set of techniques designed to help stabilize the cognitive thought process and bridge the time of cognitive degradation.

Research Design

- The study used a mixed -methods design
- First phase - survey was conducted with the the participating pilots.
  - This portion of the study was used to gauge the pilot's perceptions of their flying skills during a startle event.
  - The analysis explored for common threads of pilot thinking and reactions.
  - The results of the survey were compared and correlated to the data from the scenario sets.
- The second phase of the study involved evaluating professional airline pilots flying two different scenarios in an FAA approved level D full flight simulator (FFS).
  - The scenarios were flown by a crew consisting of a Captain and First Officer, similar to what would happen in actual line operations.
  - Each crew was presented either a low or high altitude scenario depending on the day of the week.

Startle Effect

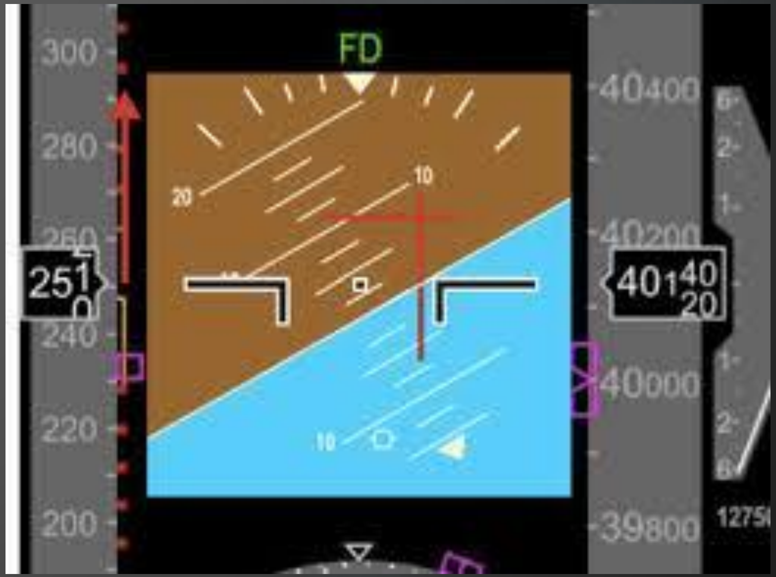
- A startle response happens when the human brain is presented with a situation that overwhelms the available cognitive resources needed to effectively handle or mitigate the situation.
- It has been widely established through psychological research that our ability to regulate our own thoughts and behaviors becomes diminished during an emotional event (Hilscher, Breiter, & Kochan, 2012)
- Research has shown that there are considerable cognitive effects on information processing following a startle event.
- The results indicate that strong cognitive and dexterous impairment could last for up to 30 seconds following a strong startle (Vlasek, 1969; Woodhead M. M., 1959; Woodhead M. , 1969; Thackray & Touchstone, 1970)

Startle Effect - Cont

- Once an unusual or emergency situation is presented, a pilot will generally be limited in their response.
- The response tends to fall into patterns a pilot has seen before, and will also be subjected to several decision-making, behavioral biases.
- An objective that is not addressed in traditional flight training is behavioral management that promotes progressive functionality under conditions of uncertainty and fear (Hilscher, Breiter, & Kochan, 2012)

Inflight Loss of Control

- Inflight loss of control is the single largest category of fatalities over the past ten years accounting for 1413 fatalities from 18 accidents (Boeing, 2012).
- Many of these inflight loss of control accidents were the result of an unusual event at the beginning of the accident sequence.
- Loss of control in flight can develop rapidly and suddenly following inappropriate decisions by the flight crew.



Research Design - Sample Groups

- The plan was to evaluate 30-60 crews (40 actual in the study). This will allow for enough data to be collected even if the effect size in the rating scale is small to medium (Cohen's d = .39 / r = .19).
  - Low Altitude – No Training (LANT)
  - Low Altitude – Training (LAT)
  - High Altitude – No Training (HANT)
  - High Altitude – Training (HAT)

Training Group

- Training consisted of a briefing and simulator practice (1hour)
- Briefing
  - Enforce a new set of callouts to identify and stabilize an undesirable situation
    - Pitch to the horizon
    - Power to 80%
    - Level the wings
    - Analyze what's wrong
    - Proper decision selection

- Basic acronym PPLA

Low Altitude Scenario

- Aircraft holding for KEWR – the hold is unexpected and the crew was released from holding to land with approximately :55 min of fuel.
  - Purpose of the hold unexpected deteriorating weather
- During the approach the crew received a landing gear warning causing a go around
- Crew performance of the event recorded at the warning
  - Scale from 1-5
    - 1 = crashed the aircraft
    - 2 = landed the aircraft in less then desirable conditions (config and fuel)
    - 3 = Landed safely with at least 30 minutes of fuel
    - 4 = situation well handled and never in doubt
    - 5 = outstanding performance.

High Altitude Scenario

Michael Gillen

- Aircraft cruising – minor deviations for convective weather
- Loss/ice-over of pitot system – loss of autoflight system
- Momentary stall indication
- Data recorded from stall indication
  - 1 = loss of the aircraft
  - 2 = excessive deviations in pitch and roll – exceeding aircraft limits
  - 3 = aircraft generally kept stable and within limits
  - 4 = good handling of aircraft with minor deviations correct diagnosis
  - 5 = excellent handling and diagnosis of problem

Evaluated Factors

	High Altitude	Low Altitude
Factors	Problem Diagnosis	Missed Approach
	Pitch	Irregular Checklists
	Roll	Time Management
	Altitude Control	Fuel Management
	Overall Control	Approach and Landing
	Overall Score	Overall Score
Seat Position	Pilot Flying	Pilot Flying
	Pilot Monitoring	Pilot Monitoring

Data Analysis

- Survey
  - Descriptive statistics
  - Lickert scale to analyze
- Data collection (SPSS)
  - Descriptive statistics
  - Linear regression
  - ANOVA to analyze training effectiveness.
  - T-tests
- Data Analysis
  - Within groups (Low and high)
  - Between groups (Total trained vs. Untrained)
  - Known standard (FAA standard ATP)

Quantitative Analysis

- A quantitative analysis was completed using SPSS on the high and low altitude profiles and the sub-factors that comprised each profile.
- The profiles were analyzed both individually and then collapsed together with the independent variable being training.
- Each of the three groups (high, low, and combined) were analyzed using a one-way ANOVA and/or t-tests.
- In addition to descriptive statistics, regression analysis was conducted on the factors that made up each individual scenario score.
- Finally, the combined group was compared to the FAA standards for Airline Transport Pilot (ATP) certification.

Data Analysis Within Group

- Analysis was conducted between Low Altitude Trained Group (LAT) and Low Altitude Non-Trained Group (LANT) and the High Altitude Trained Group (HAT) and the High Altitude Non-Trained Group (HANT).
- Analysis looked for significant findings within each group using a one-way ANOVA with the alpha level set at .05.
- Comparison between the overall grade and the FAA standard grade (3) was also compared
- A second round of analysis occurred for the contributing factors of the overall grade.
- The seat position (captain or first officer) was tested for significance as a contributor to the overall grade using regression analysis

Data Analysis – Between Groups

- The between groups were between LAT and HAT followed by LANT and HANT groups. The analysis sought to determine significant findings of the final grade, using a one-way ANOVA with an alpha level of .05.
- Linear regression was used to analyze the contributing factors and the seat position.
- Comparison to the FAA standard was also conducted in the between group comparison.

Data Analysis – Known Standard

- Collapsing across groups (all trained versus all untrained) and then comparing against a known set standard. The set standard was determined by the tolerances set forth by the FAA for an Airline Transport Pilot (ATP) certificate
- The set standard was determined by the tolerances set forth by the FAA for an Airline Transport Pilot (ATP) certificate. Crews receiving, at least, an overall grade of 3 were considered to meet the FAA standard. Any grade below 3 was considered below standard .
- The training groups (LAT and HAT) and non-training groups (LANT and HANT) were collapsed and then compared to the FAA standard using a t-test with an alpha level of .05. Comparison of the collapsed groups was also compared to the survey responses using the Pearson Correlation test

Results

- Demographics
- Survey Responses
- Flight Evaluation
  - High Altitude
  - Low Altitude
  - Low and High Combined
  - Performance verse FAA Standard
- Survey and Performance Correlations
- Results Summary

Demographics

- 40 crews who flew for a U.S. Global passenger airline participated in the study.
  - All of the subjects were active line pilots and volunteers
  - The pilots flew as a crew consisting of a Captain and First Officer and had flown in their respective aircraft for at least one year.
- Crews were also divided by which profile they flew and whether they received training and practice prior to flying the profile scenario.
  - Each scenario (low or high) was flown by 20 crews.
- In addition, crews were separated by what type of aircraft that they flew.
  - There were 21 wide-body crews (B747, B787, B777, B767) and 19 narrow-body crews (B737, A320, B757).

Quantitative Analysis - High Altitude

- Crews flew the high altitude scenario 20 times.
- There were nine untrained crews and 11 crews received the training as described in the Methods Section.
- SPSS was used to model a one-way ANOVA testing for the effects of training (independent variable) on the overall scenario event score.
- The untrained crews had a mean score of 2.67 and a standard deviation of .70, which is slightly below the standard for FAA certification (a score of 3).
- The trained crews had a mean score of 3.72 and a standard deviation of .65 which is above the FAA standard.

Quantitative Analysis – High Altitude

- A one-way ANOVA was conducted comparing the trained and untrained groups .
- There was a significant effect of training on the high altitude scenario score, $F(1, 18) = 12.25$, $p = .003$.
- The test confirms the research hypothesis that targeted training be successful in helping pilots maintain aircraft control during an unusual and sudden startle event.

Dependent Variable: High Altitude Scenario

Source	Type III Sum of Squares	df	Mean Square	F	**Sig.**	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	5.568[a]	1	5.568	12.250	**.003**	.405	12.250	.911
Intercept	202.368	1	202.368	445.210	**.000**	.961	445.210	1.000
CrewTrng	5.568	1	5.568	12.250	**.003**	.405	12.250	.911
Error	8.182	18	.455					
Total	225.000	20						
Corrected Total	13.750	19						

a. R Squared = .405 (Adjusted R Squared = .372) b. Computed using alpha = .05

Quantitative Analysis - Low Altitude

- The low altitude, low fuel scenario was flown by 20 crews.
- There were 10 trained and 10 untrained crews as described in the methods section.
- SPSS was used to analyze the results using a one-way ANOVA.
- The mean score of the untrained crews was 2.60 with a standard deviation of .70 and the trained crews was 3.70 with a standard deviation of .82.
- As with the high altitude scenario, the untrained crews performed below the ATP standards and the trained group performed above the standard.

Quantitative Analysis - Low Altitude

- A one-way ANOVA was conducted on the low altitude scenario to test the main effect of crew training.
- There was a significant effect of training on the low altitude scenario score, $F(1, 18) = 10.37$, $p = .005$.
- The test confirms the research hypothesis that targeted training was successful in helping pilots maintain aircraft control during an unusual and slow buildup of a startle event.

Dependent Variable: Low Altitude Scenario

Source	Type III Sum of Squares	df	Mean Square	F	**Sig.**	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	6.050[a]	1	6.050	10.371	**.005**	.366	10.371	.861
Intercept	198.450	1	198.450	340.200	**.000**	.950	340.200	1.000
CrewTrng	6.050	1	6.050	10.371	**.005**	.366	10.371	.861
Error	10.500	18	.583					
Total	215.000	20						
Corrected Total	16.550	19						

a. R Squared = .366 (Adjusted R Squared = .330)

b. Computed using alpha = .05

Findings – Contribuitng Factors

- Each maneuver scenario was made up of several sub-factors or components that comprised the overall score. These factors were analyzed to determine their significance in making up the total score and to uncover possible dimensions where training should be targeted.
  - The results for the high altitude scenario indicated that the most significant factor in determining scenario success was "problem identification". This was consistent with previous research findings which showed that when crews make an initial wrong decision, the in-flight issue tends to rapidly degrade.
  - The low altitude scenario was somewhat less clear in significant factors. Time management was a significant predictor of crew performance.

Findings – Pilot Flying and Aircraft Type

- The research also sought to determine if the pilot flying (Captain or First Officer) resulted in significant differences in maneuver performance.
  - The study's data demonstrated that crew performance was not affected by which pilot was flying.
  - The research question on who was flying seeks to determine if practice and/or experience could influence the overall maneuver score.
- Simulator data also did not uncover any significance between the types of aircraft in predicting the success of the scenario
  - The research question relating to aircraft type did not show it affected the maneuver scenario in either a positive or negative way. There was no significant difference in either pilot group.

Findings – Within Subjects

- In the high altitude scenario, crews were exposed to failures similar to what occurred in high altitude inflight loss of control accidents.
  - Sudden startle event
  - The group performed significantly better than the FAA standard and significantly better than the crews that flew the same profile but did not receive training.
- The low altitude scenario was modeled after several airline incidents that used increasing time pressure to induce that startle.
  - Slow startle event (building pressure)
  - The results were similar to those of the high altitude profile in that trained crews performed significantly better than the untrained crews.

Findings High / Low (Between Subjects)

- A t-test was utilized to test for significance between the mean of the trained crews for both the low and high altitude scenarios.
  - The test did not show significance in either case ($p = .89$ and $.91$)
- The untrained groups were also analyzed for significance between the low and high altitude scenarios.
  - The mean for the untrained also did not show any statistical significance with $p = .78$ and $.76$
- The results of the t-tests indicate that there was no significant difference between the trained and untrained groups when the trained and untrained groups are combined across scenario sets.
- This indicates that training was successful over a wide variety of scenarios.

FAA Standard

- The t-test revealed, that in the case of no crew training, that performance was significantly below the FAA certification standards.
- The mean score difference was -.42 that resulted in a significance of p = .01. This indicates that crew performance during a startle event is significantly different from the FAA standard.
- In the case of crew training, the results were also significant but in a positive direction resulting in a mean score difference of .71 with a p = .000.

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
No Crew Training	-3.024	18	.007	-.42105	-.7136	-.1285
Crew Training	4.564	20	.000	.71429	.3878	1.0407

Findings

- The data showed that targeted training could improve crew performance while flying startle scenarios in the simulator.
- Significance was found for the trained crews in both the low and high altitude scenarios when compared to the untrained crews.
  - Both scenarios recorded a similar main effect: power of eta = .6 suggested a medium to large effect size .
  - The effect of the training was shown to be high, predicting that trained crews would perform 73% better than untrained crews (Coe, 2002)
- Trained crews also showed a significant increase in performance when compared to the FAA standards for ATP certification.
- The results answered the research question asking whether targeted training could increase crew performance during a startle event in the simulator.

Significance

- The data recorded for this study showed that targeted training can help pilots bridge the cognitive gap when startled.
- Crews performed equally well in both the high altitude and low altitude scenario, suggesting that the training had a broad array of effectiveness.
- The study was not designed to eliminate the startle response which would be very difficult to accomplish, but sought to help crews manage the period of cognitive impairment.
- In summarizing the training, the motto "live for the next 60 seconds" was often used.

Significance – Cont.

- The training suggested in this study has implications for the airline industry as a whole.
- Following the startle event, the untrained crews lapsed out of ATP standards as described in the Methods Section of this study.
- All of the crews were eventually able to successfully recover from the simulated situation, however it is the decision making at the onset which can prove critical to event outcome.
- During the study there were no crews put the aircraft into an undesired aircraft state (UAS).
- This suggests that current airline training may be improved by incorporating startle training.
- Several published papers allude to this idea in that airline training has become rote and routine; not challenging crews with new situations and scenarios that expand flying knowledge and experience (Casner, Geven, & Willliams, 2012).

Final Thoughts

- Follow on studies should look at the effectiveness of this type of training over longer intervals.
- The training presented in this study was designed to be broad in nature and cover various states of contingencies as it relates to a startle event.
- Such training is intended to be applicable in a general way and is not intended to be aircraft or airline specific.
- Recent accidents accidents review that aircraft control was lost in the first 30 seconds following a startle event.
- Training should focus on this time period to be most effective.
- There is not a single solution in airline training to eliminate the risk of a startle event, only mitigating factors, that when presented in multiple layers serve to aid crews in successfully handling the event.

Questions

Note: The full version of this presentation and citations are available on request.

Photo used courtesy of Alexandr Vasilyev