Pilot Training Next: Modeling Skill Transfer in a Military Learning Environment

Heather M. Wojton, Project Leader
Daniel J. Porter
Emily A. Fedele
Pilot Training Next: Modeling Skill Transfer in a Military Learning Environment

Heather M. Wojton, Project Leader

Daniel J. Porter
Emily A. Fedele
Executive Summary

This product is a repackaging of content from three previously approved briefings. We have put together pieces of these briefings for a short presentation at the Human Factors & Ergonomics Society 2019 Annual Meeting. All technical content has undergone technical review and is approved by the sponsor for release.
The Pilot Crisis
Fighter pilot shortage grows to one in four, GAO says

Updated April 13, 2018
By Barrie Barber, Staff Writer
Shortage can be solved by increasing retention rate OR increasing growth rate
Retention improvements have failed

- A variety of retention efforts have been tried
  - All of these efforts have failed to solve the problem
- Lifestyle, family, and location are what matter
  - Pay bonuses aren’t enough
- If retention can’t be improved, growth must be
Traditional pipeline bandwidth is insufficient for growth

- Student quality must be maintained
  - Must attract and identify more talent

- Not enough instructors, aircraft, or money
  - Have to increase efficiency for this to work
Pilot Training Next is the exploratory attempt to overcome these challenges
Change procedures to incorporate new technology

**Industrial Age Model**
- Linear progression
- Sortie-centric
- Standardized
- Accumulates subjective data for decision

**Information Age Model**
- Emergent progression
- Task & Aptitude-centric
- Individualized
- Continual decisions based on accumulated objective data

Graphic taken from PTN-team briefing circa DEC 2017
Biggest change is making immersive sims freely available all day through COTS VR technology

Acronyms: Commercial Off-the-shelf (COTS); Virtual Reality (VR)
Simulators work to a point

- Decades of research and practice show sims work, **BUT**
  - Combat operational space is more chaotic/less constrained
  - Military hasn’t attempted the same level of reliance as civilians
  - Not all skills will transfer equally at all points of practice

- How much is enough? How much is too much?
“Learn anywhere; validate in the aircraft.”
PTN started with the T-6 to validate skill learning
PTN is on track to run one cohort per year

- Nov 2017 – PTN Stood Up
- Jan 2018 – V1 Academics
- Apr 2018 – V1 T-6 Training
- Jul 2018 – V1 T-6 Complete
- V2 Academics – Feb 2019
- V2 T-6 Start – Apr 2019
- V2 Students all Tracked – Aug 2019
- Jan 2020 – V3 Projected Start
PTN graduated their first cohort in 2018

• Started with 15 traditional students
  ➢ Pinned wings on 13 officers

• PTN accelerated and individualized the training schedule
  ➢ Compressed traditional 28-week T-6 syllabus into 8-10 weeks
  ➢ Students progressed at their own pace

• Good first pass, but not sufficient evidence to change the entire Air Force training infrastructure
PTN needs to demonstrate that skills still transfer from sim to sortie even with these new processes and technologies
Why is this hard?
Matching conditions would allow us to easily pair simulator and flight grades for modeling.
Operational realism makes skill transfer hard to model

Aileron Roll
Power-on Stall
Barrel Roll

Cloverleaf
Immelmann
Loop

G-Awareness
TP Stalls
Spin Recovery

Pitchback
Split S
Slow Flight

Aileron Roll
Spin Recovery
Lazy 8
Solution:
Compare general learning trajectories in the aircraft and simulator to evaluate transfer
Aligned learning trajectories suggest transfer

Skill transference

Less skill transference

[Graphs showing data points and trend lines for Sim Flight and Less Sim Flight]
Standard models won’t work for PTN data

- Standard modeling techniques assume:
  - Linear progression
  - Normally distributed residuals
  - Same starting point
  - Same progression for all
  - Data points are independent

PTN data violate these assumptions
Learning trajectories require special models

- Credible models of human learning must:
  - Allow people to start at different points
  - Allow people to learn at different rates
  - Account for dependencies in data

- Mixed Modeling framework is needed for human data
Building a Credible Model
Raw data are noisy
Why does performance vary between people?

- Practice
- Talent/Personal Factors
- Circumstances
- Judge/Grader
- Unknown unknowns

**Model Goal:**
Which of these matter and how much?
How do we measure amount of practice?

• Amount of formal training?
  ➢ Flight hours?
  ➢ # of maneuvers attempted?
  ➢ # of syllabus events to date?

• Self-led practice?
  ➢ Sim freeplay?
  ➢ Visualization?
  ➢ Book studying?

Days into Training
Students graded on general and specific criteria

<table>
<thead>
<tr>
<th>General</th>
<th>Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission Analysis/Products</td>
<td>Wing Takeoff</td>
</tr>
<tr>
<td>Ground Ops</td>
<td>Interval Takeoff</td>
</tr>
<tr>
<td>Takeoff</td>
<td>Instrument Trail</td>
</tr>
<tr>
<td>Departure</td>
<td>G-Warmup / Awareness</td>
</tr>
<tr>
<td>Basic Aircraft Control</td>
<td>Lead Platform</td>
</tr>
<tr>
<td>Cross-Check</td>
<td>Pitchout (Both)</td>
</tr>
<tr>
<td>Enroute Descent / Recovery</td>
<td>Fingertip (Wing)</td>
</tr>
<tr>
<td>Inflight Checks</td>
<td>Route (Wing)</td>
</tr>
<tr>
<td>Inflight Planning</td>
<td>Fighting Wing (Wing)</td>
</tr>
<tr>
<td>Clearing / Visual Lookout</td>
<td>Straight Ahead Rejoin</td>
</tr>
<tr>
<td>Communication</td>
<td>Turning Rejoin</td>
</tr>
<tr>
<td>Risk Mgmt / Decision Making</td>
<td>Overshoot</td>
</tr>
<tr>
<td>Situational Awareness</td>
<td>Echelon (Wing)</td>
</tr>
<tr>
<td>Task Management</td>
<td>Breakout (Wing)</td>
</tr>
<tr>
<td>Emergency Procedures</td>
<td>Lost Wingman (Both)</td>
</tr>
<tr>
<td>General Knowledge</td>
<td>Extended Trail (Wing)</td>
</tr>
<tr>
<td></td>
<td>Position Change</td>
</tr>
<tr>
<td></td>
<td>Formation Approach (Both)</td>
</tr>
<tr>
<td></td>
<td>Formation Landing (Both)</td>
</tr>
<tr>
<td></td>
<td>Battle Damage Check</td>
</tr>
<tr>
<td></td>
<td>Flt Integrity / Wingman Consideration</td>
</tr>
</tbody>
</table>
Defined performance as average in general category

<table>
<thead>
<tr>
<th>Task</th>
<th>NG</th>
<th>U</th>
<th>F</th>
<th>G</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission Analysis/Products</td>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ground Ops</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Takeoff</td>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Departure</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Basic Aircraft Control</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Cross-Check</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Enroute Descent / Recovery</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Inflight Checks</td>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflight Planning</td>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clearing / Visual Lookout</td>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk Mgmt / Decision Making</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Situational Awareness</td>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task Management</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Emergency Procedures</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>General Knowledge</td>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[\text{Total Score} = 2.36\]
Building learning trajectories for flights vs. sims

- Learning Trajectory Model
  - Data = all V1 T-6 syllabus events, both flights & sims

Mean General Grade =

\[
\text{Intercept} + \beta_{(\text{Simulator} \mid \text{Flight})} + \beta_{\text{Practice}} + \beta_{\text{Practice}^2} + \beta_{(\text{Practice}^2 \mid \text{Simulator} \mid \text{Flight})} + \text{Unique Intercept for each Student in (Simulator} \mid \text{Flight)} + \text{Unique } \beta_{\text{Practice}} \text{ for each Student in (Simulator} \mid \text{Flight)}
\]
Learning trajectories align in the T-6
Some FREAKs’ sim/flight trajectories looks different
Learning trajectories are suggestive, but not direct evidence of skill transfer
Step 1: Build a model using the simulator data only

Sim Data

Grade
ID
Time

Model Creation Process

Grade = ID + Time
(Sim Equation)\(^1\)

\(^1\) Equation is simplified
Building a simulator-only learning trajectory model

- Learning Trajectory Model
  - Data = all V1 T-6 graded simulator events

Mean General Grade =

Intercept
$\beta_{\text{Practice}}$ +
$\beta_{\text{Practice}^2}$ +
Unique Intercept for each Student in (Simulator | Flight)
Unique $\beta_{\text{Practice}}$ for each Student in (Simulator | Flight)
Step 2: Feed flight data to the simulator-only model

Flight Data

Grade = ID + Time  
(Sim Equation)

Predicted Grade

<table>
<thead>
<tr>
<th>ID</th>
<th>Time</th>
<th>Real Grade</th>
<th>Predicted Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>FREAK16</td>
<td>24</td>
<td>3.36</td>
<td>3.16</td>
</tr>
</tbody>
</table>
If the sim model predicts flight grades well, we have much stronger evidence of skill transfer.

Also provides insights into potential points of intervention.
Sim model fits flight data well for students who pass

- Outside 90% PI
- Within 90% PI

Did not reach MAF/CAF
Model error is twice* as high for failing students

\[ t(136.4) = -3.15, \quad p = .002 \]
PTN has provided evidence that changing processes to incorporate new technology has enabled learning transfer from the simulator to the aircraft.
Analysts can use modeled learning trajectories to evaluate skill transfer in unstructured learning environments
Lessons for other programs

• As we leverage sims more and more, it is increasingly important to evaluate how well they enable skill transfer
  ➢ Bad sims hide their acquisition cost in training

• To measure skill transfer, we need an outcome we can track over time equivalently between sim and reality
  ➢ In PTN, this was grades

• Missing data or meta-data can kill this evaluation
  ➢ High-quality data collection is more important than ever
Thank You
Backups
Air Force pipeline gets increasingly specialized

**UPT**
- Phase 1: Academics
- Phase 2: Flight
- Phase 3: Track
  - T-38
  - T-6
  - T-1

**IFF**
- T-38

**FTU**

~53 weeks

~8 weeks

Acronyms: Undergraduate Pilot Training (UPT); Introduction to Fighter Fundamentals (IFF); Formal Training Unit (FTU)
Air Force pipeline gets increasingly specialized

**UPT**

Phase 1: Academics

Phase 2: Flight

Phase 3: Track

T-6

T-1

T-38

Mobility

Bombers

**IFF**

T-38

**FTU**

T-38 availability is a major bottleneck

Acronyms: Undergraduate Pilot Training (UPT); Introduction to Fighter Fundamentals (IFF); Formal Training Unit (FTU)
PTN started with a manageable portion of the pipeline

**UPT**
- Phase 1: Academics
- Phase 2: Flight
- Phase 3: Track

PTN V1 primarily modified T-6 training

**IFF**
- T-38

**FTU**
- Mobility
- Bombers

Acronyms: Undergraduate Pilot Training (UPT); Pilot Training Next (PTN) Version 1 (V1); Introduction to Fighter Fundamentals (IFF); Formal Training Unit (FTU)
**Industrial Age Model**
- Linear progression
- Sortie-centric
- Standardized
- Accumulates subjective data for decision

**Information Age Model**
- Emergent progression
- Task & Aptitude-centric
- Individualized
- Continual decisions based on accumulated objective data

---

**Cumulative Performance**

Tasks sorted by phase and rigor

**Mastery of Cognitive & Technical Skills**

Progress Toward Completion

FTU 1
FTU 2
FTU 3
FTU 4
PTN V1 students learn through a variety of methods

**Academics**
- Instruction
- Self-Study

**Formal Events**
- Graded Sorties
- Graded Sims

**Free Play**
- Unsupervised
- “AI” Instructor
- Human IP

Acronyms: Pilot Training Next (PTN) Version 1 (V1); Instructor Pilot (IP)
The traditional models locks progression together

<table>
<thead>
<tr>
<th></th>
<th>April</th>
<th>May</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dan</td>
<td>U</td>
<td>G</td>
</tr>
<tr>
<td>Heather</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>
PTN advances students when they demonstrate proficiency

<table>
<thead>
<tr>
<th></th>
<th>Contact</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dan</td>
<td>U</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Heather</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Contact</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
</tr>
</tbody>
</table>

Acronyms: Pilot Training Next (PTN) Version 1 (V1); Unsatisfactory (U); Fair (F); Good (G); Excellent (E)
Building learning trajectories for flights vs. sims

- Learning Trajectory Model
  - Data = all V1 syllabus events, both flights & sims

Mean General Grade =

\[
\text{Intercept} + \beta_{(\text{Simulator} \mid \text{Flight})} + \beta_{\text{Practice}} + \beta_{\text{Practice}^2} + \beta_{(\text{Practice} \ast (\text{Simulator} \mid \text{Flight}))} + \beta_{(\text{Practice}^2 \ast (\text{Simulator} \mid \text{Flight}))} + \text{Unique Intercept for each Student in (Simulator} \mid \text{Flight}) + \text{Unique } \beta_{\text{Practice} \text{ for each Student in (Simulator} \mid \text{Flight})}
\]
Learning trajectories look different for flights and sims
Trajectories parallel for a period
Discrepancy occurring at the end
PTN started sims for new aircraft before V1 ended

**UPT**
- **Phase 1:** Academics
- **Phase 2:** Flight
- **Phase 3:** Track
  - T-6
  - T-1
  - T-38

**IFF**
- T-38

**FTU**
- Mobility
- Bombers

Acronyms: Undergraduate Pilot Training (UPT); Pilot Training Next (PTN) Version 1 (V1); Introduction to Fighter Fundamentals (IFF); Formal Training Unit (FTU)

Discrepancy is occurring after pilots are tracked
MAF/CAF may not be achieving transfer

But what about the T-6 syllabus alone?
Starting MAF/CAF Intro hurts sims but not flights
Pilot Training Next: Modeling Skill Transfer in a Military Learning Environment

Daniel J. Porter (OED); Heather M. Wojton (OED); Emily A. Fedele (STD);

Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

Director, Operational Test and Evaluation
1700 Defense Pentagon
Washington, DC 20301-20301-1700

Approved for public release. Distribution is unlimited.

Pilot Training Next is an exploratory investigation of new technologies and procedures to increase the efficiency of Undergraduate Pilot Training in the United States Air Force. IDA analysts present a method of quantifying skill transfer from simulators to aircraft under realistic, uncontrolled conditions.

Air Force; Pilot Training Next (PTN); Simulators; Pilot Training; Linear Mixed Models

Unlimited

Unclassified

Unclassified

Unclassified

Heather Wojton (OED)

(703) 845-6811