Defining the Training Mix – Sorties, Sims, and Distributed Mission Operations

Charles M. Colegrove
Air Combat Command DMO Team
(Alion Science and Technology)
Langley AFB VA
chuck.colegrove.ctr@langley.af.mil

Leah J. Rowe
L-3 Communications
Mesa AZ
leah.rowe@mesa.afmc.af.mil

George Alliger
The Group for Organizational Effectiveness, Inc.
Albany NY
george.alliger@groupoe.com

Michael Garrity
Aptima, Inc.
Woburn MA
mgarrity@aptima.com

Winston Bennett, Jr.
Warfighter Readiness Research Division
Human Effectiveness Directorate
711th Human Performance Wing
Mesa AZ
winston.bennett@mesa.afmc.af.mil

ABSTRACT

The fidelity and cost effectiveness of flight simulators have increased dramatically over the last decade. Given increasing fiscal constraints, many organizations have struggled with questions such as, "What is the appropriate mix of live and virtual training?", and "What is the equivalence of virtual sorties to live ones?" In July 2008, Air Combat Command (ACC) received a tasking from Headquarters Air Force to determine the optimum mix of live training, simulator training, and distributed simulator training. The best way to answer the question would involve a series of longitudinal studies for each weapon system to explore options and quantify the tradespaces in each training approach. However the time constraints imposed in the tasking prevented this time consuming and expensive approach in the near term. ACC, in conjunction with the Air Force Research Laboratory's Warfighter Readiness Research Division, developed a methodology to systematically gather data from operational personnel to help frame responses to the tasker for a number of mission areas and weapon systems and to provide a unique set of data for examining key issues in live and virtual training and the investments in and mixes of both. This paper outlines the development of the methodology, data collection instruments and techniques, subsequent analyses, and applications of the information in determining the right mix and funding profiles for live and virtual training. We will also explore the relationships between current training requirements and the outcomes of the data collection and discuss potential ways ahead and lessons learned. Plans for evaluating different mixes in operational field studies will also be discussed.

ABOUT THE AUTHORS

Charles M. Colegrove is a Senior Military Analyst for Alion Science and Technology in the Air Combat Command Distributed Mission Operations program. He has over twenty years experience in fighter operations and operations training. His main areas of focus for ACC include integrating advanced simulation capabilities into operational training programs and coordinating training research. He also leads, for ACC, the development and application of Mission Essential Competencies.

Leah J. Rowe is a Research Scientist with L-3 Communications in the Link Simulation and Training division at the Air Force Research Laboratory in Mesa Arizona. She received her M.S. in Applied Psychology in 2007 from Arizona State University and is pursuing a Ph.D. in Industrial Organizational Psychology at Capella University.

George M. Alliger is Vice President of Solutions, gOE, Inc. Dr. Alliger has extensive experience conducting performance and training projects within a wide variety of public and private organizations. He helped develop the architecture for establishing Mission Essential Competencies that has been widely adopted by the US Air Force for use with their teams. He received his doctorate in Industrial Organizational Psychology from Akron University, Ohio, in 1985.

Michael J. Garrity is Director for Aptima's Human Performance Division. His expertise is in training development, measurement, and the development of models of accident causation. Dr. Garrity leads Aptima's support of ACC and AFRL in the Mission Essential Competency program. He holds a Ph.D. in Industrial-Organizational Psychology from Clemson University.

Winston Bennett is a Senior Research Psychologist at the Air Force Research Laboratory, 711 Human Performance Wing, Human Effectiveness Directorate, Warfighter Readiness Research Division, in Mesa AZ. He is a Fellow of the American Psychological Association. He is the Technical Advisor and team leader for continuous learning and performance assessment research and development. He maintains and active involvement presence in the international research community through his work on various professional committees, his participation as a member of multinational studies, research, panels and teams, and his continued contributions to professional journals and other technical publications. He received his PhD from Texas A&M University in 1995.

Defining the Training Mix – Sorties, Sims, and Distributed Mission Operations

Charles M. Colegrove
Air Combat Command DMO Team
(Alion Science and Technology)
Langley AFB VA
chuck.colegrove.ctr@langley.af.mil

Leah J. Rowe
L-3 Communications
Mesa AZ
leah.rowe@mesa.afmc.af.mil

George Alliger
The Group for Organizational Effectiveness, Inc.
Albany NY
george.alliger@groupoe.com

Michael Garrity
Aptima, Inc.
Woburn MA
mgarrity@aptima.com

Winston Bennett, Jr.
Warfighter Readiness Research Division
Human Effectiveness Directorate
711th Human Performance Wing
Mesa AZ
winston.bennett@mesa.afmc.af.mil

INTRODUCTION

Under the leadership of Gen Richard Hawley, Air Combat Command (ACC) began a program to field and network high-fidelity simulators in 1997. The first such simulators, or Mission Training Centers (MTCs), were operational for the F-15C in 2000; other MTCs followed for the AWACS and F-16 Block 50 in subsequent years. Additionally, ACC upgraded other Combat Air Forces (CAF) simulators to provide a Distributed Mission Operations (DMO) capability. DMO is the ability to network disparate, geographically separated simulators over a world-wide dedicated network designed for persistent, daily use.

Over the years the fidelity and training capabilities of these advanced simulators have gained the attention of senior leaders in the Air Force. In July 2008 Air Combat Command (ACC) received a tasking to determine the optimum mix of live training, simulator training, and DMO.

The Air Staff's stated goal was to define "optimal" mixes. However, optimal is a bit of a misnomer within the context of the current activity. Optimality in this case is very context specific and dependent on a number of exogenous variables. Moreover, the goal is not optimal mixing, but rather the identification of the range of practical, realistic options for live and virtual mixes.

We defined the mixes that are available by mission area, weapon system type, and capacity to train with the variety of training environments that were available at the time these data were gathered and analyzed. As such, it is important to focus on the methodology that was developed, the analyses that were undertaken, the quality of the data that resulted from the methodology and the analyses, and the implications from the results for decision making regarding the range of options available by mission area and weapons system training environments.

DEFINING THE METHODOLOGY

The task fell to ACC's Flight Operations and Training Branch (ACC/A3TO), which developed a conceptual "thumbprint" diagram of the possible outcome (Figure 1).

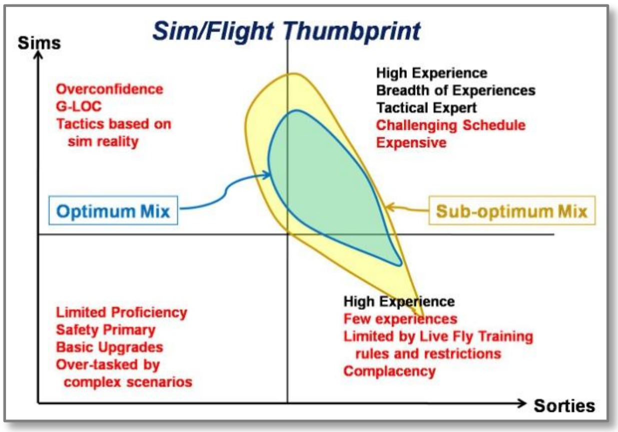

Figure 1. Conceptual Diagram of the Sim/Sortie Mix

While not initially backed by data, this basic chart specified the conceptual space of the potential relationship between the number of simulator training events (sims; y axis) and live sorties (x axis) that build aircrew proficiency. The term “thumbprint,” as applied to this quadrant chart, is simply meant to reflect the anticipated shape of the overall data – an oval with its major axis along the line formed when X+Y is equal to some empirically determined constant. This constant is expected because the potential number of sims and live fly events are presumably compensatory – an increase in one will practically limit the other.

The chart can also be meaningfully divided into quadrants based on the total of simulator and live events. Data predominantly in the upper left quadrant would represent an inordinate number of sim to live training events, presumably producing a proficiency primarily based on the fidelity of the simulators and their limited ability to replicate the real world, including the actual dynamic flight environment experienced in fighter operations.

Data that falls mostly in the lower left quadrant would signify that not enough live or simulator training events were occurring to develop combat proficiency. In the lower right, an excessive emphasis solely on live flight could result in proficiencies limited by prohibitive costs incurred by replicating complex operating environments and adversaries.

The upper right quadrant, representing large numbers of live and simulator training events, can represent ideal training. However, training providers encounter the expense limitations associated with high numbers of both complex live training environments and limits to the practical number of training events aircrew can accomplish in a given period of time.

ACC enlisted the Air Force Research Laboratory’s (AFRL) 711th Human Performance Wing, Warfighter Readiness Research Division (711 HPW/RHA) in Mesa AZ to assist in defining a methodology and tools to collect data to define the optimal live/sim region for training CAF aircrew. ACC also chose to define the mix in terms of warfighter proficiency for each CAF aircraft-based weapon system; six fighters, three bombers, and three C2ISR platforms. Although the methodology was used for all CAF weapons systems this paper concentrates on the fighter and bomber effort. Note too that CAF training taskings are recognized as a “by month” requirement. Therefore, for the purpose of this study, we used a month as the common unit of time.

What Information Should We Collect and From Whom?

Representatives of ACC, AFRL, the Group for Organizational Effectiveness, and Aptima, Inc. (the latter two organizations are consultants for AFRL) met to develop the methodology. Several factors drove the final product:

- We expected that the results would align to the concept of the “thumbprint”.
- Our choice to define the mix in terms of aircrew proficiency required a commonly recognized definition of proficiency.
- We would need to collect data from Subject Matter Experts (SMEs) in aircrew proficiency.
- ACC is the Lead Command for the CAF, so the data collection would have to provide a comprehensive answer encompassing assets and training conditions across three Major Commands: ACC, US Air Forces Europe (USAFE), and Pacific Air Forces (PACAF).

Given the above factors, we explored the possibility of using, as a starting point, the existing training specifications contained in the Ready Aircrew Program (RAP) tasking documents for each weapon system. RAP taskings are revised semi-annually and identify the number and type of training activities aircrew are required to accomplish over a period of time to maintain Combat Mission Ready (CMR) status. Most RAP taskings had, for several years, remained static in terms of total training requirements, and had not yet included new simulator training events based on the capabilities of DMO or newly acquired advanced simulation capabilities.

In the CAF, RAP is a universally recognized and used training specification. For this reason, it made sense as a frame of reference for presenting the questions to the SMEs. After considering several different combinations of the current training specifications, we decided to use the current definitions of the training requirements (i.e. Offensive Counter Air, Basic Surface Attack, etc.), but avoided indicating during measurement the actual current required numbers of live and simulator events. We did not want to imply that the current requirements would be the preferred answer.

How Do You Define Proficiency?

The CAF is responsible for providing USAF forces to Combatant Commanders (COCOMs). COCOMs through the aligned Coalition Forces Air Component Commander (CFACC) establish proficiency requirements for each weapon system based on the tasked mission types provided to the individual COCOM. The USAF has established commonly accepted definitions for proficient and highly proficient aircrews:

- *Proficient:* Squadron members have a thorough knowledge of mission area and **occasionally may make an error of omission or commission**. Aircrew are able to operate in a complex, fluid environment and are able to handle most contingencies and unusual circumstances. Proficient aircrew are prepared for mission taskings on the first sortie in theater.

- *Highly Proficient*: Squadron members have a thorough knowledge of mission area and **rarely make an error of omission or commission**. Aircrew are able to operate in a complex, fluid environment and are able to handle most contingencies and unusual circumstances. Highly proficient aircrew are prepared for the mission taskings on the first sortie in theater.

Who are the SMEs and Where Do You Find Them?

Field commanders, Operations Group and Squadron Commanders, are ultimately responsible for ensuring aircrew under their command are properly trained to accomplish assigned missions. Daily training is overseen at the squadron level and administered and assessed by unit instructors – the same personnel that would also be at the core of combat operation execution. Unit commanders, directors of operations, and instructors became our target SME group.

While RAP taskings are constant among the three commands, training conditions and local requirements have some minor variations. To refine the optimum mix answer we had to gather data from the above mentioned SMEs across the CAF – 25 bases stretching from the Western Pacific to Europe.

THUMBPRINT DEVELOPMENT WORKSHOP

In order to identify estimates for the number of live and simulator missions necessary to develop and maintain proficiency at the levels previously defined, we needed to create a method to arrive at these estimates reliably. During the initial meeting at AFRL, we designed a sample matrix based on the current definitions for training missions assigned to each weapon system. The intent of the matrix was to solicit expert opinions with respect to the number of specific training events, by type, over a given period of time, required to maintain the previously described proficiency levels.

After the initial method was developed it was necessary to test the method in order to identify shortfalls, to clarify directions, and to understand the potential/impact of different methods between communities (i.e., the B-2 community may use a slightly different method to arrive at estimates for live/fly mix as compared to the F-16). We hosted a workshop during which we asked the Functional Area Managers (FAMs) for the F-15E, F-22, and B-2 to refine definitions and develop a methodology. That methodology would be used by experts in their community to identify an optimal mix of live/sim training events, tailored to the need of their specific community. FAMs are the ACC staff officers from the individual weapon system communities and are responsible for developing and coordinating RAP taskings for their community.

After an introduction and orientation from the ACC project manager, AFRL presented information relating to competency-based measurement and each FAM received relevant Mission Essential Competency (Colegrove and Alliger, 2002) information for their specific platform. As pre-work, the study designers asked each FAM to refamiliarize themselves with specific COCOM requirements and current RAP taskings.

We then asked the FAMs to complete several tasks:

1) How well can a given RAP mission be addressed via Live Fly, via Sim? 1=Not at all, 2=Slightly, 3=Moderately, 4=Substantially, 5=Completely

2) Develop estimates for the amount of training necessary to meet COCOM requirements in four areas: Simulation (Standalone simulator period and DMO) and Live (Local flying training and Large Force Employment)
3) By aircraft, identify a time unit of analysis for estimating number of events (1 month, 5 months, 20 months) for experienced and inexperienced aircrew (Sirkin, 2005)
4) Lastly, sum the individual columns to obtain the ratios between live and sim training

The groundwork performed by the FAMs became the baseline for the refinement of the final survey matrix described later in this paper. During the workshop, we noted many similarities and differences between communities. A summary is provided below:

F-15E
- Used a modified definition of experienced as 500 hours (400 in the aircraft plus 100 in the simulator)
- 1 DMO event per month due to availability
  - Assumed the simulator would be available
- Training events include planning, execution, and debrief (for both live and sim events)

B-2
- Used the current book definitions of experience (500 aircraft hours).
  - Assume that the remaining (non B-2 aircraft or sim) time per month is flying T-38s
- The B-2 will use DMO to integrate with other Global Strike assets
- Training events include planning, execution, and debrief

F-22A
- Used the book definition for experience (500 aircraft hours)
- Multi-level Security for DMO is in place

The Thumbprint Workshop set the groundwork for developing the Thumbprint Surveys as well as the operational definitions that would be required to carry out the effort. In the following sections we will discuss the testing and fielding of the survey.

THUMBPRINT SURVEY TESTING AND FIELDING

Operational Definitions

The previously discussed definitions for Proficient and Highly Proficient are commonly accepted in the Air Force. However, in order to collect meaningful data it is critical that all respondents have a common understanding of the key concepts being assessed. The following training environment definitions were also given to each responder:

- Training Event: Live or simulator events where you brief, fly, and debrief
- Stand-alone simulator: Simulators not connected to another unit's sim. Multiple unit simulators may be connected to form a flight (e.g. four F-15s or two B-1s that are locally networked).
- DMO: Simulation with the capability to link to another unit's sim.
- Local Live Fly Training: Local training at your home unit.
- Local Composite Force Training (CFTR)/Large Force Employment (LFE): Scenarios employing multiple flights of aircraft, each under the direction of its own flight leader acting in a LFE scenario to achieve a common tactical objective. Scenarios should be opposed by air and surface threats and should include at least 8 blue aircraft.

Thumbprint Survey Structure

The Thumbprint survey was divided into four different sections. Section one gathered some preliminary demographical information such as Mission Design Series (MDS) hours and rank. Section one also asked the following question: “What is the maximum number of training events that you can realistically accomplish in one month?” This question was used to gauge the feasibility of the responses in sections two and three and also provided an indication of aircrew availability for training.

In section two the participants were asked to, “Enter the number of training events that YOU FEEL are required to be **PROFICIENT** across a 6 month period of time in the above mentioned four environments for an inexperienced and experienced pilot”. This section provided the data to calculate the optimal training mix for inexperienced and experienced pilots to be proficient at their taskings.

Like in the previous section, section three asked the participants to, “Enter the number of training events that YOU FEEL are required to be **HIGHLY PROFICIENT** across a six month period of time in the four environments for an inexperienced and experienced pilot.”

Finally section four included several subjective value-added questions. The purpose of this section was to assess a number of areas where the quality of the data and the capacity to capture realistic data depended upon the point of view and the understanding of the responders to the data collection activity. Moreover, there were a few questions that did not fit neatly into a standard Likert scale type of assessment or data collection methodology. Therefore, a number of specific, though open-ended questions were developed to help further "unpack" the rater's frame of reference and to obtain data on several areas of relevance. Those questions follow:

- If you were part of a team completely redesigning your training program that included live fly and high-fidelity simulation, what decision would you make with regard to the ratio of simulator training to live fly training? Please select the mix that would be optimal. (Respondents were given the options of more, same, and less for each simulator and live environments). Please provide rationale for this decision.
- What type of simulator does your unit have access to?
- What shortfalls and training gaps do you see in your unit's simulator?
- Fiscal realities may dictate that some training be migrated to a high-fidelity simulator. If you had unlimited access to a high-fidelity simulator training environment what training can only be achieved during live flight?
- What training can only be achieved in the simulator?
- Please prioritize the order of importance of having live flight training in each of the following missions (1=most important and X=least important) (the respondents were given the list of their respective RAP events).
- What are the benefits and/or disadvantages of relying on a regional sim training center (TDY required) vs. an MTC at home station?

Later, during the analyses it turned out that the data from these questions were instrumental in both explaining the foundation from which the ratings in the other parts of the surveys came and for identifying key information about the items that the field responders felt were only applicable to either simulation or live training. These then served as frames for the remainder of the data analyses and discussions for that mission area, weapons system, and relevant training environments.

Once the general structure of the survey was developed pilot testing began to assess the validity and viability of the survey.

Pilot Testing

During the Thumbprint Workshop a general methodology was defined to measure the optimal number of training events that are required for an experienced and inexperienced pilot to be proficient and highly proficient. In order to assess this tool four phases of pilot testing took place. The primary concern was nailing down a time frame that the responders would rate. At the time of the study, the RAP training cycle took place during a twenty month period.

During Phase I of pilot testing we assessed the methodology using a 20 month period. In order to test the survey, three recently retired F-16 SMEs completed the survey. The SMEs indicated that rating across a twenty month period was too burdensome and difficult to conceptualize. They recommended that a three month time frame be used instead.

During Phase II of the pilot testing the survey was revised to a three month span. This survey was pilot tested with three different SMEs (retired F-16 pilots). Contrary to Phase I it was recommended that a three month time frame was not enough time to complete the ratings because there are certain tasks or missions that may not be completed as frequently as every three months (e.g. LFEs). Following the survey assessment an informal interview between the SMEs and the researchers took place where the SMEs indicated that a squadron typically forecasts training using a six month schedule and that our target sample would most likely relate best to this time frame.

Phase III of the pilot testing utilized operational pilots as upon the completion of the first two phases of this study we were confident that the six month time frame was sufficient. The responders consisted of ten instructor-qualified Air National Guard F-16 pilots who were participating in a training research exercise at the AFRL in Mesa AZ. The results of this pilot test indicated that the measure had face validity and yielded consistent results.

During Phase IV of pilot testing the survey was administered to the rest of the ACC FAMs from each MDS. The FAMs were able to provide refined input to the process. A portion of the survey is at Figure 2.

			Inexperienced <500 flight Hours						Experienced ≥500 Flight Hours					
		Month →	1	2	3	4	5	6	1	2	3	4	5	6
DCA	SIM	Stand-alone simulator												
		DMO												
	Live Fly	Local live fly training												
		Live CFTR/LFE												
		Total per month												
OCA	SIM	Stand-alone simulator												
		DMO												
	Live Fly	Local live fly training												
		Live CFTR/LFE												
		Total per month												
TI	SIM	Stand-alone simulator												
		DMO												
	Live Fly	Local live fly training												
		Live CFTR/LFE												
		Total per month												

Figure 2. Thumbprint Survey Rating Matrix

Data Collection

The Thumbprint survey was administered to operational aircrew over a time span of three months. Data was collected at twenty operational units from 317 fighter and bomber responders. The survey was proctored at each site by at least one of the Thumbprint team members.

DATA ANALYSIS

The primary focus of this effort was to capture the optimal ratio of live to simulator training events per month. Therefore, following the data collection effort the surveys were analyzed and charted. Recall that each responder rated the number of training events required to be proficient and highly proficient for an experienced and inexperienced individual. These ratings provide four different data points:

1) Inexperienced Proficient
2) Experienced Proficient
3) Inexperienced Highly Proficient
4) Experienced Highly Proficient

For the rest of this paper we will used a hypothetical case as an example. The purpose of this paper is to highlight the methodology rather than the results.

Outlier Analysis

An outlier analysis was completed prior to analyzing the data. The importance of a conservative outlier criteria was recognized due to this being data provided by experts. Therefore, we selected three standard deviations from the mean as the outlier criteria allowing us to retain as much data as possible in the sample while recognizing the necessity to exclude extraneous data. As a result any responder's data that was three standard deviations from the mean or greater were eliminated from the analyses (Winer, Brown, & Michels, 1991).

Ratio "Optimal Training Mix" Analysis

In order to obtain the ratio of the "optimal training mix" of live to simulator training per month the data provided by sections two and three of the survey were analyzed.

First, a mean for each training environment live and simulator was calculated for each responder. Therefore, the two simulator environments were combined as were the two live environments. This data provided the number of training events required per month in each of the following conditions:

1) Inexperienced Live Proficient
2) Inexperienced Simulator Proficient
3) Inexperienced Live Highly Proficient
4) Inexperienced Simulator Highly Proficient
5) Experienced Live Proficient
6) Experienced Simulator Proficient
7) Experienced Live Highly Proficient
8) Experienced Simulator Highly Proficient

The data from these eight conditions represents the number of events that each responder feels is required to meet these conditions.

The next step was to calculate the optimal live to simulator mix ratio in the four categories of interest (i.e. inexperienced proficient, experienced proficient, inexperienced highly proficient, experienced highly proficient). For each condition listed above, the mean of each responder's answer was calculated. Accordingly, we calculated the mean across each MDS which provided a single number for each of the eight conditions listed above. An example of the ratios can be found in Table 1.

Table 1. Hypothetical Number of Live to Simulator Training Events Required per Month "Thumbprint Optimal Mix"

	Proficient		Highly Proficient	
	Live	Sim	Live	Sim
Inexperienced	15	7	21	8
Experienced	12	5	17	7

In order to report the data in a meaningful manner they were presented in charts and split based on level of proficiency. Rather than only showing the optimal mix we choose to show the array of data that represents the individual data points as well. These charts provide the viewer an opportunity to gain a thorough understanding of the Thumbprint survey results. An example of a Thumbprint chart is shown in Figure 3.

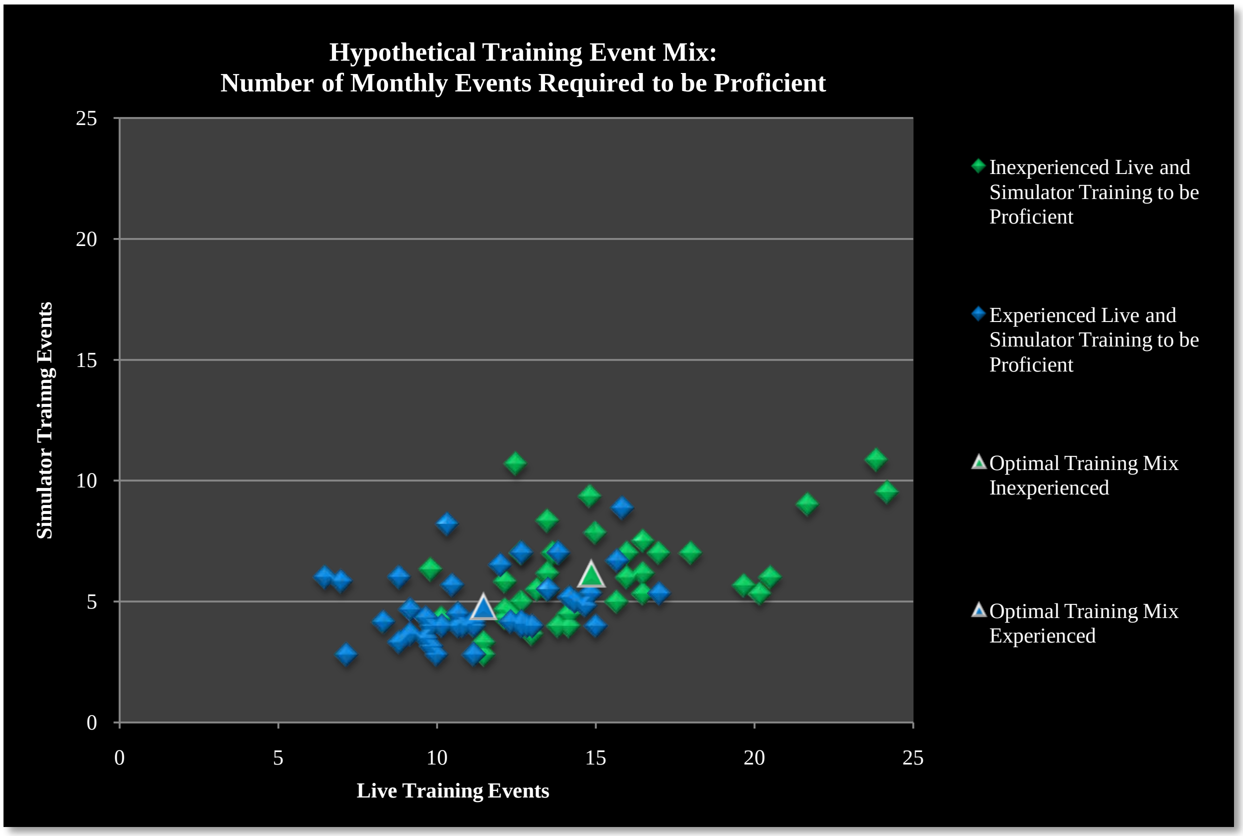

Figure 3. Hypothetical Proficient Thumbprint Chart

Instrument Validity and Reliability

One important question in looking at any set of data is to understand its degree of validity and reliability. The validity of the Thumbprint survey results lies in the concept of content validity – namely, that a rigorous process of instrument construction and data collection can help ensure that a process (in this case, a survey) measures what is intended. An assertion of content validity does not preclude the desirability of other ways of examining validity (e.g., convergent, discriminant, or construct); however it does act as a serious foundational claim.

The estimation of the reliability of survey data can also be approached in different ways. Fundamental to the concept of reliability is the idea of replicability, whether across time or similar entities. It is in the latter sense that we gauged the reliability of the "thumbprint" surveys. Specifically, we computed what are termed Intraclass Correlation Coefficients, or ICCs. An ICC can be understood as a correlation reflecting degree of agreement among judges of some phenomenon. Again, in this specific case, the judges were the SMEs who completed the surveys, while the phenomenon being judged was the number of times in a month that it was considered desirable for an event to occur, as explained earlier in this paper. Figure 4 below is an example of ICCs computed for a single platform, the F-15C. Note that most of the ICCs are above .80 – this is a threshold above which interrater agreement can be characterized as "excellent." These same indices were computed for a substantial sampling of the other platforms, and results were similar.

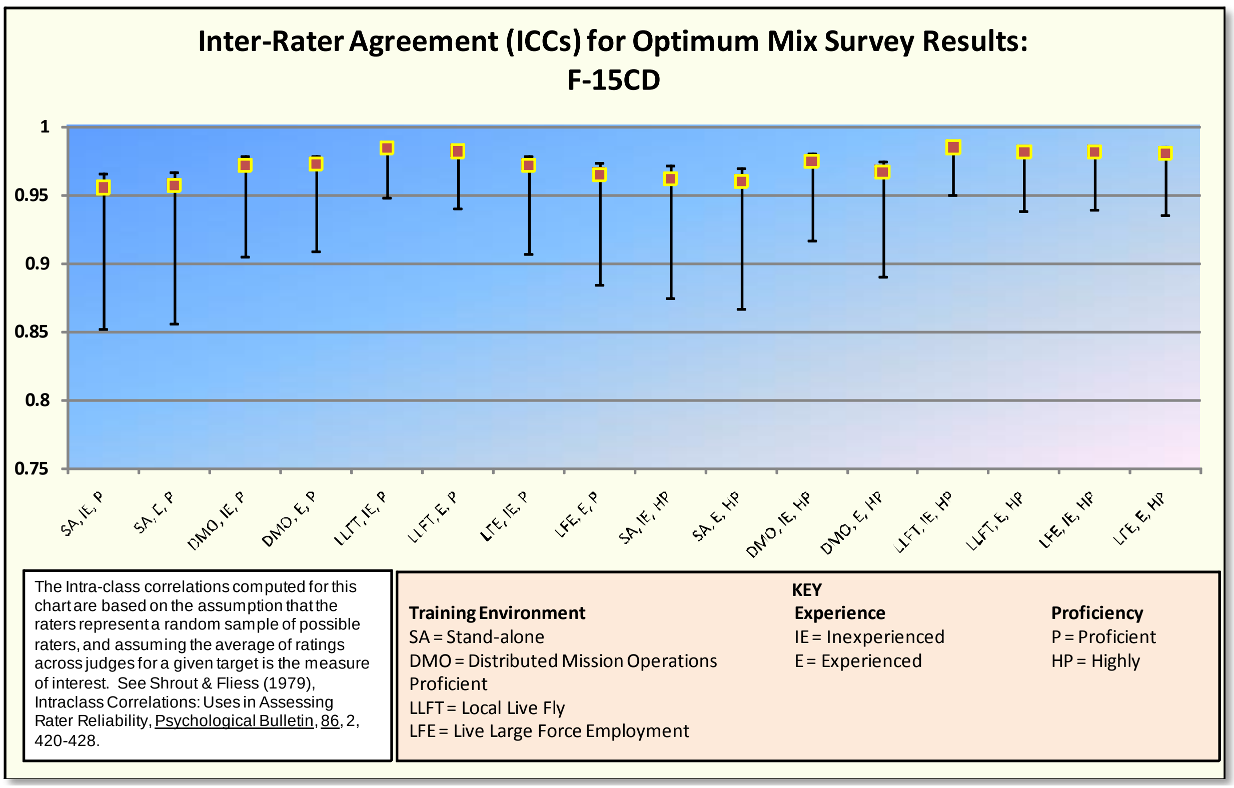

Figure 4. Sample Intraclass Correlation Coefficients

APPLYING THE DATA

The discussion below references the notional data in Table 1 and is focused on the number of events for the inexperienced aircrew to be proficient. For our example the numbers of events are fifteen live sorties and seven simulator missions.

Aircrew Availability

Section one of the survey addresses the maximum number of training events that aircrew can practically accomplish in a month. While this number varied between the fighter and bomber communities they compare favorably given that live bomber missions usually, because of the sortie duration, require more than one day to accomplish the required mission preparation, execution, and debriefing. Fighter missions are normally accomplished in one day. Aircrew availability that exceeds required/desired events offers the opportunity to increase training requirements however, availability that falls short may require further studies aimed at increasing availability and is not within the scope of this paper.

The Best Media for the Event

In section four of the survey we asked if they would prefer, in a revised RAP tasking, more, the same, or less live and simulator training requirements, and the rationale behind that answer. In the same section we also asked which mission types or events could be done only live or only in the simulator. When combined, these responses provided insight into internal ACC questions (and subsequent Air Staff questions) related to the training that can be accomplished solely in the aircraft or in the simulator. Generally, the SMEs pointed to complex environments and force organizations being more readily available in the virtual environment. Dynamic flight (decision making under "G") and mission elements that require accurate visual representations of the real world were common responses for missions and events that can only be trained in the aircraft.

The Cost of Training

Resultant Thumbprint charts, when compared to the existing training requirements, offer insight into the effectiveness of the current training specification as well as an indication of the potential cost to train at the optimal level. In addition to the cost to operate aircraft flown by the training audience, the cost of live training also includes the training infrastructure (e.g., emitters, targets, airborne adversaries, and range support) and can easily be an order of magnitude higher than the cost to operate simulators and associated constructive environments. Large virtual exercises can also have costs pertaining to multiple environment generators and a supporting white force for exercise management that, while not in the scope of this paper, still provide a less costly training alternative to large live exercises.

One of the purposes of the study was to determine areas for investment and potential costs associated with that investment. Current training requirements that exceed the optimum mix do not point out the need for investment in the training system. The rest of this section is devoted to an optimal mix that exceeds current training requirements, ways to increase training opportunities, and investment.

Additionally, notional RAP requirements of eleven live sorties and four simulator missions per month are used in the comparison. In Figure 5 below, optimal mix data is plotted as the green triangle and RAP requirements are the blue "plus".

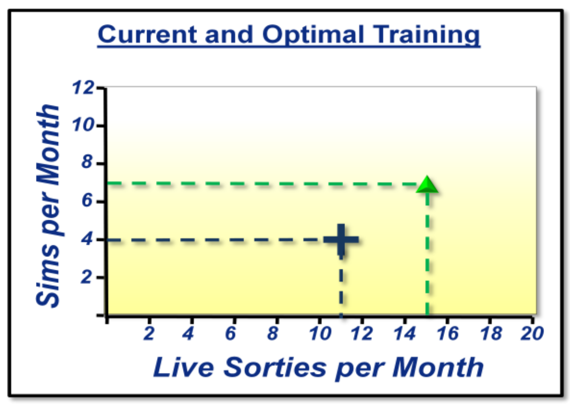

Figure 5. Comparison Between Optimal Mix and RAP Requirements

For the discussion below we have used the notional values for the thumbprint data and costs of $10,000 per live flying hour and $1000 per simulator hour.

Table 2 compares the cost of the current "state of training" with the cost to train at the optimal rate determined by the SMEs in our example.

The costs in Table 2 are per pilot/crew per month. It is important to note – the total figures are meant to identify a steady state training cost (i.e., the cost per month to train the number of live and simulator missions required by each pilot or aircrew) and not associated infrastructure costs.

Table 2. Cost Comparison Between the Current State of Training and Training at the Optimum Mix

	Current	Optimal Mix
Live Sorties/Mo	11	15
Cost/Live Sortie	$10,000	$10,000
Sim Missions/Mo	4	7
Cost/Sim Mission	$1,000	$1,000
Total Training Cost/Mo	$114,000	$157,000

Making Up the Difference – Determining Investment Strategies

In our example the optimal mix exceeds the current RAP requirement by seven total events (four live and three sim) and the difference in cost is $43,000 per month. However, there are other considerations when attempting to increase the number of required training events that may require analysis of the investments required developing additional infrastructure and strategies to provide funding. Any exploration of increasing training requirements must consider aircrew availability to participate.

A major consideration is the capacity of the infrastructure to support increased training. To meet the optimal mix, increasing simulator training over the current capacity may be limited by operating hours and staffing. This option may require additional investment to provide more training opportunities within an existing device. Impacts to aircrew duty day requirements must be evaluated when considering extending simulator operating hours.

DMO, as discussed at the beginning of this paper, is the ability to link disparate, geographically separated simulators to conduct warfighter training. DMO may be considered an addition or expansion to a required sim mission and therefore aircrew availability is influenced by the same factors discussed above. The additional consideration for increasing DMO participation is the availability of the dedicated DMO infrastructure to support an increased demand.

Another, potentially more expensive, option to meet an increased training requirement is the acquisition of additional simulators and staff at a given location. However, the potential success of this option also hinges on aircrew ability to man an increased number of simulators at the same time.

Increasing live operations has the same considerations with respect to aircrew availability but adds important infrastructure questions. Can maintenance support an increased sortie production? Are ranges adequate (in size and capacity) to support more sorties? Do adversaries (both ground-based and airborne) exist in sufficient numbers to support the level of training?

CONCLUSIONS

Dangerous Assumptions

Optimal, appropriate, proper, and achievable are all adjectives that are being applied to combinations of live and simulator training. However, there may be a tendency to orient the end result to the least expensive solution. ACC chose to define the problem and resulting answers based on the end product – warfighter proficiency.

Overly optimistic reports of simulator capability and well publicized demonstrations may raise the expectations of those that are not involved in the daily realities of training. These expectations may include the idea that there is a direct one-for-one trade between simulator training and live training. This study did not show that to be true.

There were identifiable ratios (approximately two live sorties per sim mission for units with high-fidelity sims; 3:1 for medium-fidelity; and 4:1 for low-fidelity) for a given total optimal mix. Given the time constraints of the tasker and the volume of survey work necessary to compile the data we did not conduct excursions from the totals provided by the SMEs.

Positive Results

As a result of this large field survey effort, fighter and bomber training managers are implementing new, and usually increased, sim training requirements. At the end of the current revision cycle sim training will comprise between 20 and 30% of the required training within the CAF.

AKNOWLEDGEMENTS

Special thanks to Lt Col Chris “MOTO” Davis, ACC/A3TO Branch Chief at the time of this writing. He is responsible for the “thumbprint” concept and graphic at Figure 1.

REFERENCES

Colegrove, C. and Alliger, G. (2002). Mission Essential Competencies: Defining Combat Mission Readiness in a Novel Way, *SAS-038 Symposium, Brussels Belgium*

Shrout & Fliess (1979). Intraclass Correlations: Uses in Assessing Rater Reliability, Psychological Bulletin, 86, 2, 420-428.

Sirkin, R.M. (2005). Statistics for the Social Sciences, 3rd Edition. Thousand Oaks, CA:Sage.

Winer, B.J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.