

# **ACTION RESEARCH: A STATISTICAL REPORT ON IP UPGRADE**

**Gary Heartsill, PhD**

6 January 2010  
(Completed 22 Aug 2015)

*Three Distinguished Graduates*



*A sample of or our training*

*“In action research, it is the teacher who identifies the Research topic..., the teacher who collects information to investigate the topic, and the teacher who intercepts and judges the research results..., The teacher is at the center of action research.”*

(Gay & Airasian, 2000, p. 293).

**ABSTRACT**

**Teachers need to improve their practice by engaging in professional growth and action research suggests using data to inform how well the students perform. A questionnaire survey has been used in the C-21A IP Upgrade class for most of the 750 graduates. This report reflects some of the results from the selected sample data of 400 students who were able to take the Personality Profile, which was incorporated in the middle of 2002. This draft report reflects the descriptive statistical findings and compares the data with a previous survey of 100 students in 2000. Highlighted are the student class evaluation scores demonstrated both in a one-sample *t* test and used as the dependent variable in multiple linear regression with the rest of the survey equation variables. The one-sample *t* test was significant,  $t(399) = 21.82, p < .01$ .**

*History*

The Air Force C-21A Instructor Pilot Upgrade class<sup>1</sup> has been under my teaching supervision since January 1998 and during this time about 754 students have attended the class; indeed, this period has been the most fulfilling teaching assignment I have had the privilege of working in. So, for all of us educators (instructor pilot candidates included) who care about developing criteria to identify and narrow the specific teaching aspects of this practice, here is my brief report.

*Action Research*

Working with a couple of Gay and Airasian’s beliefs that underlie action research the questionnaire can help identify these beliefs: (a) “Teachers should have opportunity to engage in professional growth, (b) Teachers want to improve their practice and need data to do so, and (c) ...Teachers are able to carry out action research studies that will inform their practice” (p. 593).

Some of the research questions include (but are not limited to):

1. How much improvement is there between the pre tests and posttest scores and is there any statistical significance in the results?

---

<sup>1</sup> The term Fundamentals of Instruction (FOI) is another name we operate under so this FOI may show up later on in the text. This draft was completed on 6 January 2010 and reformatted on 22 Aug 2015.

2. How do the Air Force Academy graduates compare with the civilian college graduates?
3. How does age effect (or correlate) with the other variables?
4. How does flying time or active duty time correlate with the other variables?
5. How well do the female pilots do with the males?
6. To what degree does the personality profile fit into the correlation of the students?
7. To what extent do the high scores in the Student Evaluation correlate to the other variables (if any) and is there any significance in the results?

### *Questionnaire*

Each class has been given a questionnaire that includes demographics, a test on their feelings toward instructing, and a test on their feelings or attitude toward flying. A pre test is given on the first day and a posttest is administered on the last day followed by a student evaluation of the five-day academic training. In the fifth year a Personality Profile (perpro.exe) was administered to determine the students profile from one of nine choices.

The questionnaire variables are:

1. Student class FOI number
2. Last four numbers of social
3. Gender (1 – male, 2 – female)
4. Age
5. College (1 – AFA, 2 – Civil, 3 – other)
6. Active Duty – in years
7. Flying Time
8. Attitude toward Instructing
9. Attitude toward Flying
10. Personality Profile (there are nine)
11. Pretest score
12. Post test score
13. Evaluation

Four hundred students (N = 400) were chosen as a sample of the population for this study. For sure, an n number of about 200 would have been more than enough to use but I wanted to use the numbers and data for as large a group as I could manage for a couple of reasons (Gay, Mills, & Airasian, p. 111).

- More students will be able to see their personal results from the questionnaire.
- Using 400 personality profiles will give more depth.
- Using 400 personality profiles will better show statistical dispersion.
- “By increasing the sample size, the standard error is decreased and the power of the test is increased” (Hinkle, Wiersma, & Jurs, 1988, p. 305).<sup>2</sup>

---

<sup>2</sup> The argument in descriptive research is to use as large a sample as possible. I have been waiting for this number of 400 for a long time as it now will give me a better chance to be able to generalize the findings to the population - and maybe the whole Air Force. “Beyond a certain point (about N = 5,000), the population size is almost irrelevant and a sample size of 400 will be adequate” (Gay & Airasian, p. 135).  
For sure, there is a question of randomness in this present method and I am aware of this possible violation.

An example of case lines (student lines) entered into SPSS (vs. 13.0, 2004) are:

```
001 310 5446 2 26 2 4 800 80 57 1 60 85 96
002 310 8874 1 28 2 4 1200 74 64 2 50 65 91
.
.
280 632 8683 1 28 2 3 800 84 79 2 35 75 100
```

### *Results*

The descriptive statistics shown below in Table 1 reflect the mean scores for most of the variables and of interest for the practice is to see only 6% of the candidates are female, 40% graduated from the Air Force Academy (3% from other military academies), the average age of 28.3 – a rather young age, average flying time is just over 1300 hours.

The test scores for the pre and posttest show an improvement after the treatment of the five days and, although not shown in this paper, the results are statistically significant and, for sure, the direction a teacher should hope the grades go.

**Table 1**

*Results of the Descriptive Statistics (Frequencies)*

---

Variable	<i>n</i>	<i>M</i>	<i>SD</i>	%
Gender 1	377			94
Gender 2	23			6
Age	400	28.3	3.6	
Col 1	154			40
Col 2	237			59
Col 3	3			1
Active Duty		5.8		
Fly Time		1335		
ATTI	400	87.34	8.06	
ATTF	400	75.0	7.14	
Pretest	392	50.92	10.98	
Posttest	392	70.08	10.98	
EVAL	399	95.80	4.39	

---

*Note:* The Personality Profiles are expanded and displayed in Table 2.

The personality profiles are just shown to show the descriptive results as the procedures and algorithms for further study have not been completed yet. The research question number six is under more study. See Table 2 for the distribution.

**Table 2 Personality Profiles**

**PERSON**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Perfectionist	105	26.3	26.5	26.5
	Helper/Caretaker	68	17.0	17.2	43.7
	Achiever/Performer	12	3.0	3.0	46.7
	Authentic Individualist	11	2.8	2.8	49.5
	Observer/Philosopher	46	11.5	11.6	61.1
	Loyal Team Member	9	2.3	2.3	63.4
	Optimistic Generalist	47	11.8	11.9	75.3
	Powerful Individualist	73	18.3	18.4	93.7
	Team Player	25	6.3	6.3	100.0
	Total	396	99.0	100.0	
Missing	999	4	1.0		
Total		400	100.0		

The results of the two surveys are shown in Table 3 reflecting very similar scores. More research will be completed in this area to see if some particular scores correlate.

**Table 3**

*Results of the Descriptive Statistics (Frequencies) compared to the 2000 survey.*

Variable	<i>n</i>	<i>M</i>	<i>SD</i>	%	2000 <i>n</i>	<i>M</i>	<i>SD</i>	%
Gender 1	377			94	92			92
Gender 2	23			6	8			8
Age	400	28.3	3.6		100	29.9		
Col 1	154			40	53			53
Col 2	237			59	47			47
Col 3	3			1				
Active Duty		5.8			100	7.8		
Fly Time		1335			100	2000		
ATTI	400	87.34	8.06		100	88.95	6.55	
ATTF	400	75.0	7.14		100	73.3	6.57	
Pretest	392	50.92	10.98					

Posttest	392	70.08	10.98			
EVAL	399	95.80	4.39	100	90.76	6.47

*Evaluation Scores and results*

Research question number 8 asks “To what extent do the high scores in the Student Evaluation correlate to the other variables (if any) and is there any significance in the results?” Regression was run on most of the variables and the only one that showed any correlation was attitude toward flying (ATTF) and it was not much. The tests were designed to meet the learning objectives and do not have the normal test taking attributes nor do they have the reliability and validity “normal” tests are supposed to have. Having said that, more research will be given to how to improve the examinations, as the training objectives in the community appear to be changing.

However, due to the high scores given in the evaluation variable some time was spent to see what some of the reasons were and to, again in question 8, see if there was any significance. The data suggests the scores have improved over the first 100 scores given in 2000. This is shown in the equation as the Test Value = 91. The results listed below show the results as significant and the effect size as a large effect. See Table 4.

Of course, having a high positive student grade response on the class evaluation<sup>3</sup> is a good indication – in most cases - of training and teaching going in the right direction. More time will be spent on this area perhaps in researching the comment section to see how the specific training scores can be improved.

**Table 4**

**T-Test**

**One-Sample Statistics**

	N	Mean	Std. Deviation	Std. Error Mean
EVAL	399	95.80	4.395	.220

**One-Sample Test**

	Test Value = 91					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
EVAL	21.823	398	.000	4.802	4.37	5.23

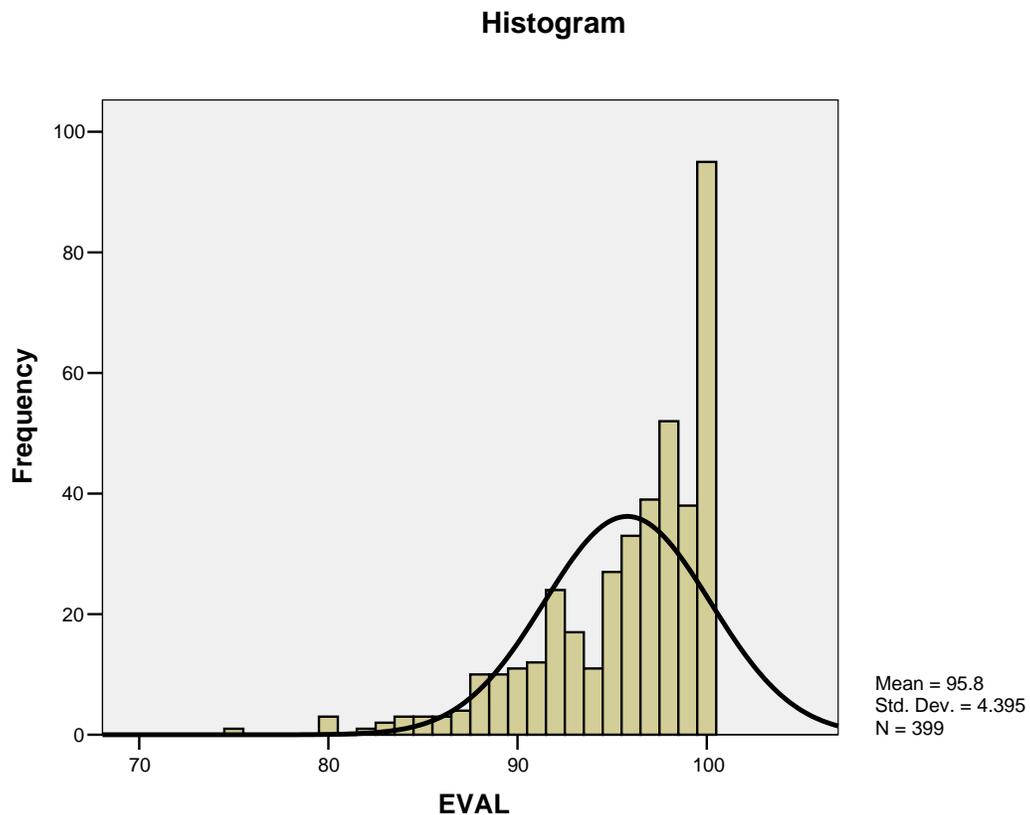
A one-sample *t* test was conducted on the EVAL scores to evaluate whether their mean was significantly different from 91 (90.76),  $t(399) = 21.823$ ,  $p = .000$ . The 95%

<sup>3</sup> Although not given in this survey, under the rubric of summative evaluation, guidance is given in our AF Manuals, particularly in 36-2236 (2003).

confidence interval for the EVAL mean ranged from 4.37 to 5.23. The effect size  $d$  1.09 indicates a large effect (Green & Salkind, 2005). Table 5 shows the distribution of EVAL scores. The observed value of  $t$  is 21.823 is greater than the critical value of 1.960 (Gay, 2000, p. 615), so the null hypothesis of the means being the same as the test value is rejected. Therefore, the test is significant,  $t(399) = 21.82, p < .01$ .

See Table 5 for the graphic description of the EVAL scores. See Appendix A for additional descriptive statistics.

**Table 5**



*Conclusions (truncated)*

The questionnaire has shown areas of strength and areas needing some shoring up to meet the cognitive issues of training and learning to be an instructor pilot.

## REFERENCES

- Air Force Manual 36-2236. (2003). *Guidebook for Air Force instructors*. Government Printing Office.
- Coram, R. (2002). *Boyd: The fighter pilot who changed the art of war*. New York, NY: Little, Brown and Company.
- Gay, L. R., & Airasian, P. (2000). *Educational research: Competencies for analysis and application*. Upper Saddle River, NJ: Prentice Hall.
- Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data*. Upper Saddle River, NY: Pearson Education.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1988). *Applied statistics for the behavioral sciences*. Boston, MA: Houghton Mifflin Company.
- SPSS. (2004). SPSS vs. 13.0. Chicago, IL.

## APPENDIX

The picture below, during a class field trip to the third floor, demonstrates the issues of John Boyd's Energy Management (EM) theory (Coram, 2002, p. 147).



**Ps [T – D/ W] V**

## APPENDIX A

### Evaluation Descriptive Frequency Statistics

EVAL

N	Valid	399
	Missing	1
Mean		95.80
Std. Deviation		4.395
Skewness		-1.373
Std. Error of Skewness		.122
Percentiles	25	93.00
	50	97.00
	75	99.00

EVAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	75	1	.3	.3	.3
	80	3	.8	.8	1.0
	82	1	.3	.3	1.3
	83	2	.5	.5	1.8
	84	3	.8	.8	2.5
	85	3	.8	.8	3.3
	86	3	.8	.8	4.0
	87	4	1.0	1.0	5.0
	88	10	2.5	2.5	7.5
	89	10	2.5	2.5	10.0
	90	11	2.8	2.8	12.8
	91	12	3.0	3.0	15.8
	92	24	6.0	6.0	21.8
	93	17	4.3	4.3	26.1
	94	11	2.8	2.8	28.8
	95	27	6.8	6.8	35.6
	96	33	8.3	8.3	43.9
	97	39	9.8	9.8	53.6
	98	52	13.0	13.0	66.7
	99	38	9.5	9.5	76.2
100	95	23.8	23.8	100.0	
Total		399	99.8	100.0	
Missing	999	1	.3		
Total		400	100.0		