Analysis of the Decision Making Process of Flight Instructors at the Brazilian Air Force Academy (AFA)

Luiz Maurício de Andrade da Silva, Laércio Aparecido Lucas, Maurício Pereira da Costa

Brazilian Air Force Academy

lma28@uol.com.br

ABSTRACT

Motivation – Discusses the human judgment in decision making, through the analysis of the flight instructor's accuracy at the Brazilian Air Force Academy, when attributing grades to the cadets in flight training. **Research approach** - It is possible to increase elements to the studies of the Naturalistic Decision Making confronting the attributed grades in flight with those obtained by the cadets in an aptitude test for the military pilotage. **Research limitations/Implications** - This is ongoing work. A deeper exam of the psychological questions would possibly bring new insights to this study. **Originality/Value** - Taking grades obtained by the cadets in the aptitude test, a regression model was generated and it served as base for the verification of the accuracy of the instructors' judgment. This analysis means a positive referential in the validation of the NDM, as well as their unfolding in the aviation, the Aeronautical Decision Making - ADM. **Take away message** - The results of this work are elucidating in subjects of interface man-computers.

Keywords:

Decision making – human judgment – aeronautical decision making

INTRODUCTION

The practice of flying and - maybe primarily - the teaching of flying techniques are some of the most fertile situations to the non controlled study of the decision-making process, particularly the human judgment when making decisions.

In the ambit of Brazilian Air Force (FAB), flying lessons are realized at the Air Force Academy (AFA), where, besides learning the practice of flying (in the second and last year of the course), the cadets receive education in Administration, and also in technical/specialized courses of military doctrine.

The necessity to ally excellence in flight instruction with the optimization of the process, prevails there, seeking for more efficiency in the use of the available resources, once the equipment involved (airplanes), as well as its operation (fuel, technology support and maintenance) show high costs and availability restrictions.

One of the factors that compete with the objectives of optimization and efficiency in flight instruction is the high evasion observed during the teaching process, after the cadets' first contact with the experience of pilotage.

This evasion – that can occur after significant theoretical investment has already been done in the cadet's formation – often is associated to the fact that some students simply do not have the minimum aptitude necessary for piloting, or at least not to fly as a military pilot.

There is, however, the necessity to undertake an effort in the sense to detect the aptitude (or not) before the beginning of the practice of piloting, substantially reducing the volume of resources employed.

This is the way that computer programs of flight evaluations were developed to serve as auxiliary instrument in selecting novice students, based on previous diagnose of aptitude. AFA obtained the English software Pilapt (Pilot Aptitude Tester).

The intention of this article is to present the results of preliminary studies that have already began in the ambit of AFA, so as we can follow the effectiveness of using these test, and their actual contribution to the process of teaching pilotage to cadets, specifically the decreasing of evasion. The decreasing of evasion in coherent basis certainly will depend on the acuity in judging and decision making process of flight instructors, who are, at last analysis, who determine the evasion or the continuity of a cadet in formation.

Aiming yet to contribute to the studies of *Naturalistic Decision Making* (NDM, according to Klein, 2000), and also to highlight new ideas about the famous trade-off between velocity and acuity in the decision making process, and, once we are analyzing the human judgment in a dynamic context, we intend with this work to offer more elements for the necessary sedimentation of NDM (Lipshitz et al, 2001).

Through the grades obtained by the cadets in the Pilapt software, we will generate a mathematical model, based on multiple regression, that will become the basis for acuity verification of the grades attributed by the instructors during flying lessons.

This verification, with (i) the grades obtained in the test, and (ii) those attributed by the instructors, in flight, will permit to realize an analysis of the acuity of the human judgment (Hastie & Dawes, 2001) realized by the flight instructors from AFA.

We understand that this analysis – besides being important in optimizing the efforts at AFA – perhaps it could mean one more referential to validate this important and innovatory theoretical tendency represented by the NDM studies.

The next section will show a review of literature related to the topics of interest in this article, followed by the methodology for data analysis and its interpretation. At the end, our conclusions and some suggestions for future studies.

THEORETICAL FUNDAMENTS

We revisited some topics in the literature about decision making process, for those we believe the preliminaries results of research presented here – and their conclusions – could be valuable.

With respect to the classical approach of decision-making process this paper contributed with nothing. To more behavioral and cognitive approaches (Tversky & Kahnemann, 1992; Hastie & Dawes, 2001; Gigerenzer, 2002; Bazerman, 2006) maybe our propositions will shed some light, even in an already well fertile environment.

To the so called organizational decision making (ODM, Danne & Pratt, 2007), *lato sensu* and *naturalistic decision making* (NDM, Klein, 2000), in particular, we expect that some of our propositions and conclusions, will proficuously illuminate, specifically with the fertilization of debates about the intuition question in the human judgment.

That said, we will start to narrow the focus of the present literature review, in the aspects regarding ODM and NDM, to then, at the end of this section, present some details about the Pilapt software.

Organizational decision making (ODM)

While the great decisions, in accordance to the so called Scientific Administration, by Frederick W. Taylor and Henri Ford, restrict to how much to produce, and to which level of standards to do it, these same decisions could be based only and exclusively in the assumption of rationality. Rationality comprehended as a closed system, with well know variables and deterministic behavior. Even though the emphasis by Henri Fayol to the role of managers would not change significantly the rational view of the decision making process throughout the XIX century.

On the contrary, before Herbert A. Simon's studies, specifically his boundary of rationality, the studies of Taylor, Ford, Fayol, Weber and Etzioni would only reinforce the impetus of a rational vision that had its apex with the quantitative theory of F. Lawrence and T. A. Edson.

Only after the brightness from Simon's prepositions, succeed it was by the marvelous reinforces offered by the systems theories from Lyon Bertalanfy, besides the theory that advocate the inseparability of the duo structure and strategy, by Alfred Chandler, and even – why not include – the thesis of inseparability between system and environment proposed by Maturana and Varella (1995), through its concept of *autopoiesis*, that one starts to realize the ODM assuming a less positivist direction.

This less positivist orientation will then find in the binomial intuition (Dane & Pratt, 2007) and subjective probability (Tversky & Kahneman, 1992; Vianna, 1989; Silva, 2000), its highest expression. Regardless it would not dispel of the classical trade-off enunciated in the beginning of this article, that is, acuity versus velocity in decision.

Subjective Probability

Since Bernoulli, in the beginning of XVIII century, passing by Vianna (1982), many authors (Kyburg & Smokler, 1965 cited by, 1982, 1989) has considered the subjective probability as the degree of belief and the state of knowledge of an individual.

The Gordian knot is to get to a number that not only reflects this degree of belief, but also be reliable under the prism of logic and of some existing postulates. So that it is not all or any degree of belief that could be accepted as subjective probability. The subjective probability as it should be interpreted as the state of knowledge of one individual.

The literature present options to adjust or calibrate some situations in which striking discrepancies are diagnosed among the subjective opinions and the relative frequencies observed (historicals) to the random variable studied. The calibration is made by training, once, still according to Vianna (1982:7), "*the person whose distribution a priori is of interest, usually is not familiarized with probability and statistics and, as a result, it is very difficult for her to speak in terms of probability, for more intuitive that the notion of probability could appear to the statistician*"

Tversky and Kahneman (1992), as well as Tversky and Koehler (1992), authors originated from psychology, have studied the existing sources of bias when resorting to the estimative of subjective probability: to the motivational order and to the cognitive order. In the motivational order, the individual consciously or unconsciously blurs his opinion, for

example, as cited by the authors, to the “satisfaction of a third one”. This, analyzed in the light of decision making in organizations, can be very complicated if we consider, among others, the coalitions of power knowingly existent in organizations (Bertero, 1989).

Yet the bias of cognitive order is connected to the individual way of judgment that can, according to the cited authors, be of five different types: (a) availability, (b) adjustment and fixation, (c) representativeness, (d) non-declared premises, and (e) coherence.

In (a) the individual resorts to experiences that have had in similar situations, and here one must have special attention to the fact that the individual, naturally, will give more attention to the more recent and easier to memorize informations. Vianna (1982:8) suggests that, once identified these situations, it makes that the individual “seek to envision landscapes of occurrence of extreme values of the random variable whose probabilities are been estimated”

In (b) the individual tends to fixate in only one datum that was initially given or that he estimated by himself. The disposition, non-sequential, of values from the random variable, for him to estimate probabilities, is demonstrated by the cited authors as a way to eliminate the bias.

In(c) the individual takes as representative very small samples from the studied variables. In (d) the individual presents a certain degree of unattachment, not feeling responsible for foreseeing extreme situations that could alter his estimates, like wars or economic crises among others.

In (e) the individual bias his opinions based on situations or indirect landscapes the he believes can come to occur and interfere in the studied variables.

Vianna (1982) scrupulously discussed aspects related to the estimative of subjectives probalibilities, from only one individual, and from juries.

The same author (Vianna, 1989: p.42) still discussed the axioms of subjective probability, and highlighted that for complex systems specific methods as Delphi are recommended. The premises pointed by Vianna (1989: p.46), for the cited method, were resumed as:

1. It is correct to request opinion from specialists when there is intention to estimate future probabilities;
2. The information obtained by groups will have higher aggregated value if compared to those obtained individually;
3. The level of individual disinformation is higher than the group one;
4. Meetings face-to-face tend to produce bias;
5. Anonymous groups are more efficient in this type of work and permit correction of bias.

AFA adopts as a decision criteria regarding the permanency or not of a cadet in their ranks, a collegiate called Counsel that will always give the final word in this question. Still, another possibility that maybe could came to be considered in the ambit of AFA is to adopt, in the referred Counsel, or at the teaching department level (Squadron of Flight Instruction – EIA), previous instance to the Counsel, and where the first signal of evasion appears, after the flight instruction, some of the techniques of decision making by panel Delphi, as cited above.

Intuition.

Erik Dane and Michael Pratt (2007: p.33) define intuition as a “judgment that appears through quick, unconscious and holistic associations”. Yet Tversky & Kahneman (1992) defines intuition as being the “thoughts and the preferences that come to mind rapidly and without much deliberation”.

Other authors (Gigerenzer, 2002; Klein, 2000) understand that the intuition to be valorized should be that which results in adequate decisions, according examples and experiences that can be extracted from observation of the work from some specialists (experts), such as firemen, doctors and airplane pilots, among others.

Bazerman (2006) refers to two types of thoughts that we use to judge and decide: one, less structured, more intuitive, fast and less based on calculation, known as system 1. And other more structured, based on calculations, therefore slower, known as system 2.

One very promising way bearing to clarify the famous trade-off seams to be the conceptualization of “intuitive judgment” (Dane & Pratt, 2007), that would consider not only the adopted process by those who should utilize them, but still its quality and its specific results. And that, furthermore, evidently would be classified as belonging to system 1.

We understand that this article can contribute to the propositions by Dane & Pratt (2007: p.43-47), in this theoretical alley of “intuitive judgment”, once we, when analyzing the specific results of the judgment realized by the Air Force Academy flying instructors (see the items methodology of research and interpretation of results), will shed lights, less over the process but more over the quality of decisions and their specific results.

Naturalistic decision making

In a brief historic about the NDM, Lipshitz et al, (2001) emphasize as necessary new studies for that the NDM comes to consolidate in a more definitive way in the field of decision making process.

Among the studies suggested by the cited authors, based on conferences about the subject, prior to 2001, is can be highlighted: (i) the importance of time pressure over the decision to be taken, (ii) the necessity to increase the studies on how the experts decide, and (iii) another trade-off, this time between the formation of the big picture versus the decision results.

In relation to the first highlight of NDM, in other words, the importance of time pressure over the decision to be taken, the question is found at AFA in a latent way and under two aspects. Firstly the pressure of the cadets' formation time: the flight solo should happen after an average of 14 hours of instruction. Secondly, and perhaps more important than the former, comes the fact that the exercise of piloting an airplane is eminently a practice in which the pressure of time decision is permanently, given the acuity demanded in the procedures as a way to reduce the risk of the operation.

For the second highlight given above, the necessity of amplify the studies about how the experts decide, we go back to what was already enunciated in the beginning of this text, as being the mainly contribution expected from this work.

Regarding the third highlight, we understand that the trade-off between the formation of the big picture versus the decision results, in this analysis we will be privileging a better comprehension of both variables, once the flight instructors of AFA should be developing some very particular criteria of combination of the instantaneous judgment of each particular case that arrives for instruction, without however, being careless with the consideration about the future consequences of their decisions.

Still according to the same authors, the mainly objective of NDM would be "to understand how people make decisions, in real life contexts, and that are familiar and significant for them." (Lipshitz et al, 2001, p.332).

This implies we think of a class of researches not controlled in laboratories, but that be involved in its natural environment. For that, the authors recommend a specific attention (from researches in NDM) to two major aspects, that are, the fact that the cognitive process is the most important, and not the technical attributes of the decision maker, and secondly, the fact that the orientation that the experts usually adopt is more orientated to a good framing of the question to be decided, and less to a previous estimation of its consequences.

We hope the results in this article come to elucidate exactly these questions.

THE PILOT APTITUDE TESTER (PiIApT)

The Pilot Aptitude Tester (Pilapt) system was developed by the Psytech LTD and PeoTech, and it had the first studies realized in England by the *Royal Air Force*.

The system has tests whose objective is the evaluation of individuals in several areas related to the psychology of aviation, comprising in a psycho-diagnostic computerized system.

The Pilapt contains a total of 6 (six) tests (Deviation Indicator, Trax, Hands, Patterns, Concentration and Capacity) that are developed directly in a computerized system controlled by a mainly server. The system has as objective to measure the yield of diverse elements that are related to ability of piloting.

The general result of Pilapt is obtained through the sum of the tests' results: Deviation Indicator, Trax, Hands, Patterns and Concentration. The Capacity test generates an individual result that is not computed to the general result of Pilapt and, for this study, it will not be utilized, given its nature of combining execution of three tasks simultaneously.

The motor coordination is investigated in the instrument in a similar way to the one produced by a pilot that is receiving instructions to accomplish a flight. The measured variables are: numerical and verbal reasoning, visual perception, visual and psychomotor coordination, special processing, memorization and use of information, processing of audible information and conversion to visual information, selective attention, and capacity of decision.

Once each one of the tests of the system will enter in our regression as independent variables, we understand to be appropriated a more detailed explanation of each test.

The tests

Deviation Indicator: Compensatory test of bi-dimensional tracking that require answers by joystick. It is related to the basic ability to control the aircraft.

Trax: It is a test of tri-dimensional tracking persecution, that besides measuring coordination, it also demands from the candidate association with special parameters related to actual flight. This test demands answers only through joystick.

Hands: It demands that the candidate processes an auditory message sent by headphone and then conducts a visual searching of the information showed on screen. The answers by keyboard are timed and the test evaluates information preceding from special abilities and orientation.

Patterns: It is essentially a test of fitted figures providing a measure of focused perception (the ability to ignore distractions and identify relevant information). Candidates should answer the visual information in a restrictive window of time and then should complete each question submitted to the stress of time. This is related to the selective attention and the ability to make decisions in an emergency situation.

Concentration: It is a dynamic measure of attention. In the first task the candidate indicates combinations of shapes and colors inside a matrix where the columns are defined by colors and the rows are defined by geometric shapes. During the test, the colors and shapes defined by rows and columns change. The candidate should click in the correct combination before it disappears using the mouse.

The dependent variable of our regression will be the grade obtained by the cadets while in flight. This grade, called "Flight Grade", is explained below.

Flight Grade: Arithmetic mean of all flights that comprehend the phases of Pre-Solo (PS), Maneuvers and Acrobatics (MAC), flight in Formation (FR) and Navigation, all referring to the Basic Flight Instructions Stage. They receive degrees going from 1 (dangerous) to 6 (excellent), in each completed flight.

METHODOLOGY OF RESEARCH

The research accomplished in this article was based on two basic criteria. The first (first criterion) is a simple direct comparison between the results from the aptitude test Pilapt, and the proportion of cadets that were approved (success) in the flying instruction against those who were not approved (fail). The second (second criterion) is a multiple linear regression that, obtained by the aptitude data of previously executed tests, allowed us to make an inference about the probable grades to be obtained in a real instruction. This inference was later confronted with the grades actually attributed by the flight instructors.

We understand that these criteria, specially the regression, allow us to evaluate the quality of human judgment when deciding, in a non-controlled natural environment, bringing to light new evidences to the NDM.

INTERPRETATION OF THE RESEARCH RESULTS

We are going to interpret the results of this research through the previously proposed two criteria sequence.

First criterion: direct comparison of stens

Following below is the relation between the Pilapt result, in stens (scale going from 1, lowest level, to 10, highest level) and the proportion of successes and failures in the basic flight instruction taught in the 2° Squadron of Flight Instruction (2° EIA) in the year of 2006 (n = 155) and 2007 (n = 135), in a total of 290 cadets. These cadets undertook instruction in the referred year with an approximated total of 44 hours of flight instruction, having undertaken the test before beginning the flight instruction activity.

For this, the cadets were divided in thirds through the stens of grades obtained from the execution of Pilapt: 1° = stens from 1 to 3 corresponding to the level of lower potential to the learning of flight instruction; 2° = stens from 4 to 7 corresponding to the level of medium potential to the learning of flight instruction; 3° = stens from 8 to 10 corresponding to the level of high potential to the learning of flight instruction

In this way, we have the following quantity of cadets distributed in each third: 1° = 35; 2° = 194; 3° = 61; in a total of 290 cadets.

For that we have the table below with the respective percentages:

Thirds	Successes	Failures	Total
1° (1-3)	20 (9%)	15 (24%)	35 (12%)
2° (4-7)	151 (66%)	43 (68%)	194 (67%)
3° (8-10)	56 (25%)	5 (8%)	61 (21%)
total	227 (100%)	63 (100%)	290 (100%)

Table 1. Stens, successes and failures

The low benefit in the flight instruction and, in some cases, failures from the course happens many times due to difficulties in adaptation to the flight activity; difficulties to manage the flight and to have dynamic attention; anxiety in higher level in relation to the instruction and learning from the cadets; difficulty of spatial reasoning; lack of motivation for the flight activity, among others. These factors are difficult to be diagnosed before the beginning of flight activity

and show themselves, for the most part, only with the beginning of flight though reports of flying instructions and from the cadets themselves, as verified by the Squadron Psychologist that supervenes the Flight Instruction.

These factors previously related are not diagnostic by Pilapt (primarily the anxiety in the face of evaluation and performance in the flight activity, and the motivation for the flight activity) or by many other psychological tests. However, it is notorious that as higher the result of Pilapt (stens), lower it is the chance of the cadet to be disconnected in flight for not showing potential to learning in the flight instruction, and as lower the result of Pilapt, higher are the chances of the cadet come to be disconnected from the instruction.

Following are tables showing a general view of what happened in the flight instruction during the years 2006 and 2007 (n=290):

Table 2: Expectation Graph

In total, from the 290 cadets that started the flight instruction, 63 (22%) were disconnected of the stage for not showing minimum conditions for military piloting. We can verify that 80% of the cadets that obtained sten 1 in Pilapt were disconnected from the flight activity, while all the cadets that obtained sten 10 concluded the course successfully.

The percentage of success in the flight instruction decrease with the reduction of stens obtained by the cadets.

Second criteria: multiple linear regressions

The regression was realized using the grades from the Squadron of Flight Instruction – Flight Grade (which varies from 1 to 6) as dependent variable (using the grade obtained in the instruction with a total of 399 cadets) and the results from the tests of Pilapt as independent variables (deviation indicator, trax, hands, patterns and concentration). The results were compared with a total of 113 Aviator Cadets that had finished the Flight Instruction in the year of 2007, obtaining the following results:

Regression	R ²	Signif. (95%)	Error
Flight Grade	0.13	3.23 x 10 ⁻¹¹	2.0%
PS	0.093	2.76 x 10 ⁻⁷	4.3%
MAC	0.064	8.45 x 10 ⁻⁵	2.3%
FR	0.14	4.02 x 10 ⁻¹²	1.1%

Table 2. Multiple Linear Regression

The Flight Grade of the Squadron has an average of 4.26 and a standard deviation of 0.30. The grade from the prediction has an average of 4.25 and a standard deviation of 0.09. The correlation between the squadron grade and the prediction grade is 0.52 with level of significance at 0.01, showing a high correlation between the two variables.

Transforming the difference between variables in percentage, we have an average of 1% and standard deviation of 26%, showing a standard rather close between the grade attributed by the squadron and the prediction grade, although the r-square was 37% of the result was not by chance.

CONCLUSIONS

As we saw, the research here discussed was presented by two criteria: a direct comparison between the stens obtained by the cadets at the Pilapt test versus the success and failure cases in instruction (first criteria), and a multiple linear regression that utilized the grades attributed by the instructors to the cadets, in flight instruction, as dependent variable; and the grades of Pilapt as independent variable (second criteria).

It can be concluded based on the first criteria, namely the direct comparison between stens and the success and failure cases, to be evident that as higher the result of Pilapt (stens), lower is the chance of the cadet to be disconnected in flight for not showing potential for the flight instruction, and as lower the result of Pilapt, higher are the chances of the cadet to come to be disconnected from the instruction.

Yet in what concerns the second criteria, as can be seen in Table 3, the observed levels of error between the grades of flight instructors and the estimative from multiple linear regression were very low, indicating quality of judgment by the instructors, as well as efficiency by the Pilapt software.

Albeit the low levels obtained by the coefficient of determination, it can be observed, by the level of significance for a degree of confidence of 95%, that the expected parameters in the regression were completely acceptable.

The question of false positives and false negatives continues to be open, that is, cases where the estimates were success in instruction and then resulted in failure, or vice-versus. In these cases it is likely that a deeper exam of the psychological questions – from the cadets as well as the instructors – would possibly bring new insights in the effort to improve flight instruction at AFA.

REFERENCES

- Bazerman, Max H. (2006). *Judgment in managerial decision making*. Sixth edition. Hoboken: John Wiley & Sons.
- Bertero, Carlos Osmar. (1989) *Cultura organizacional e instrumentalização do poder*. In: Fleury, M. T. L. & Fischer, R. M. (org.). *Cultura e poder nas organizações*. São Paulo: Atlas.
- Burke, E., & Valsler, C. (1999). *The Pilot Aptitude Tester (PILAPT) Handbook*. People Technologies.
- Dane, Erik & Pratt, Michael G. (2007). *Exploring intuition and its role in managerial decision making*. *Academy of management review*. Vol. 32, No. 1, 33-54.
- Gigerenzer, Gerd (2002). *Reckoning with risk: learning to live with uncertainty*. London: Penguin Books.
- Hastie, Reid & Dawes, Robyn M. (2001). *Rational choice in an uncertain world: the psychology of judgment and decision making*. Thousand Oaks: Sage.
- Johnston, Neil, McDonald, Nick, Fuller, Ray (1994). *Aviation psychology in practice*. Aldershot: Ashgate, 1994.
- Kaempf, George L. & Klein, Gary (1994). *Aeronautical decision making: the next generation*. In: Johnston, Neil; McDonald, Nick and Fuller, Ray. *Aviation psychology in practice*. Aldershot: Ashgate.
- Klein, Gary (2000). *Sources of power: how people make decisions*. Cambridge: MIT Press.
- Lipshitz, R., Klein, G., Orasanu, J., Salas, E. (2001) *Taking stock of naturalistic decision making*. *Journal of Behavioral Decision Making*. Vol. 14 No. 5, 331-352.
- Maturana, Humberto & Varela, Francisco (1995). *A árvore do conhecimento*. São Paulo: Wokshopsy.
- Silva, Luiz Maurício de Andrade da (2000). *Instrumentalização do planejamento estratégico: aplicação no setor aeroviário comercial brasileiro*. Tese de doutoramento. FEA/USP.
- Silva, Luiz Maurício de Andrade da (2004). *Tomada de decisões em pequenas empresas*. São Paulo: Cobra.
- Tversky, Amos & Kahneman, Daniel (1992). "Advances in prospect theory: cumulative representation of uncertainty." *Journal of risk and uncertainty*, vol.5.
- Tversky, Amos & Koehler, D. J. (1992). *Support theory: a nonextensional representation of subjective probability*. *Psychological Review*.

- Vianna, Nadia W. H. (1982). Probabilidade subjetiva e o júri de especialistas. Dissertação de mestrado. EAESP/FGV.
- Vianna, Nadia W. H. (1989). A subjetividade no processo de previsão. Tese de doutoramento. FEA/USP