

Análise Estatística da Aplicação da Escala Cooper-Harper na Avaliação da Qualidade de Pilotagem de uma Aeronave

Alvimar de Lucena Costa Junior

Instituto de Pesquisas e Ensaios em Voo - Praça Mal. Eduardo Gomes, 50 - CEP: 12228-901 – São José dos Campos – SP

Ricardo Suterio, Dr.Eng.

Instituto Nacional de Pesquisas Espaciais – Av Astronautas, 1758 – CEP 12227-010 – São José dos Campos - SP

Resumo — Este artigo relata a análise de significância de graus da escala de qualidades de pilotagem de Cooper-Harper atribuídos por seis Pilotos de Ensaio em Voo em sete Tarefas de Pilotagem, utilizando-se a técnica da Análise de Variância e da análise dos resíduos dessa distribuição. O estudo constata que os graus atribuídos por cada Piloto de Ensaio diferem pouco entre si, trazendo evidências de que a aplicação da Escala Cooper-Harper atribui maior objetividade a uma avaliação inerentemente subjetiva.

Palavras-Chave — Ensaios em Voo, Qualidades de Pilotagem (QDP), Escala Cooper-Harper.

I. INTRODUÇÃO

Durante o projeto e desenvolvimento de uma aeronave, uma das fases mais difíceis para os engenheiros de ensaios em voo é a avaliação das qualidades de pilotagem. Para um engenheiro, a aeronave é um grande sistema dinâmico, com vários subsistemas, que responde previsivelmente a um conjunto de inputs, de acordo com o que ele determinou em seus cálculos, códigos e planilhas. Porém, quando os inputs idealizados do engenheiro são substituídos pelos comandos de um ser humano no controle da aeronave, um novo sistema dinâmico se revela, bem diferente do que foi desenvolvido na tela do computador do engenheiro.

Por outro lado, para um piloto, a aeronave é um equipamento único, do qual ele espera mesma resposta padrão para cada mesmo comando inserido. Na realidade, cada aeronave reage a comandos de uma maneira específica, podendo esta resposta variar inclusive entre aeronaves de um mesmo modelo em variantes diferentes, o que pode resultar em diferentes graus de controlabilidade. Esta variação de controlabilidade pode ser desejável ou não.

As duas visões unilaterais são avaliadas durante um processo de desenvolvimento de aeronave por meio de avaliações de equipagens de Pilotos e Engenheiros de Ensaios em Voo. Estas equipagens devem traduzir as dificuldades de controle do avião ou do helicóptero de forma tão objetiva quanto possível, permitindo às equipes de engenheiros projetistas corrigir partes do sistema que geram respostas que contrariem o senso comum para um piloto.

Lucena A.L.C.J., lucenaalcj@ipev.cta.br, Tel. +55-12-3947-7958; Suterio R., suterio@lit.inpe.br, Tel. +55-12- 3208-6270

O método mais amplamente utilizado e referenciado para que estas avaliações sejam padronizadas é o uso de escalas padronizadas, sendo a mais difundida a Escala Cooper-Harper (CH), apresentada formalmente por G. E. Cooper e R. P. Harper em 1966 [1].

O método com o uso da Escala CH parte do desenvolvimento de tarefas de pilotagem (TP) por Pilotos (PP) e Engenheiros de Ensaios em Voo (EP). Uma TP, na definição de uma avaliação de QDP é definida como o “trabalho real” designado a um piloto. A TP representa um segmento de uma ação específica que compõe as fases de um voo, seja um segmento das ações para uma decolagem, um ataque ao solo, ou para uma aproximação por instrumentos. [2]

A facilidade de cumprimento de cada tarefa é avaliada por meio da escala CH, que define as QDP da aeronave ensaiada com notas de 1 (aeronave melhor controlável) a 10 (aeronave incontroleável). A Fig.1 mostra os graus CH e seus significados.

Além dos graus CH atribuídos pelo piloto ao desempenho da aeronave no cumprimento de tarefas de pilotagem, cada tarefa testada tem os comentários qualitativos do piloto catalogados, de forma a complementar e enriquecer o conceito resultante da nota atribuída pelo Piloto de Ensaio em Voo (PP).

Para que este resultado seja tão objetivo quanto possível, durante sua formação, o PP é qualificado para ser capaz de traduzir suas sensações e opiniões sobre a controlabilidade e facilidade de pilotar de uma aeronave segundo os graus da escala CH. Em geral, um piloto, desde sua formação, é treinado a aprender a voar uma aeronave, não importando quais dificuldades vá enfrentar, suprimindo uma avaliação da QDP instintivamente. O PP é treinado a perceber exatamente o contrário, e ter condições de avaliar, objetivamente, que a aeronave tem problemas de controle, não o piloto.

Em paralelo, o EP é qualificado para extrair de um piloto descrições que possam ser enquadradas da maneira mais fiel e objetiva possível a estes graus CH. O EP também desafia o PP, de forma que a dupla tenha certeza de que o grau CH atribuído pelo piloto está em consonância com os comentários subjetivos deste. O EP também analisa os dados numéricos da instrumentação do voo, que indiquem quão suave ou controlável a aeronave objetivamente foi, e também

usa estes dados para desafiar ou confirmar os graus CH do PP.

Dessa forma, o PP, nesta fase dos ensaios, pode ser comparado a um medidor calibrado para aferir o grau CH de uma aeronave para várias tarefas de pilotagem, que segmentam as várias fases de voo de uma aeronave. A referência [2] traz informações mais detalhadas a respeito das definições desta escala Cooper-Harper. A referência [3] traz maiores informações sobre o emprego e a operacionalização do uso desta escala.

Por fim, a campanha de ensaios de uma aeronave, normalmente, implica na utilização de um grupo de PP, cada um dando graus CH para as mesmas tarefas, de maneira a formar uma amostragem de graus oriundos de diferentes avaliadores, com diferentes backgrounds de experiência e conhecimento. Tal uso intensivo da cara mão de obra de equipagens de ensaio em voo justifica-se no fato de que os avaliadores não têm plena confiança nas avaliações de um único Piloto de Ensaios em Voo, considerando que o mesmo poderá trazer um *bias* à avaliação caso seja o único a dar grau CH a uma tarefa de pilotagem específica.

Para este trabalho, foram analisados os graus CH atribuídos por seis PP, para mesmas TP executadas em um mesmo modelo de aeronave durante uma campanha real de ensaios em voo. Por meio da sequência de análise estatística apresentada por Montgomery em [4], foram efetuadas análises de variância nos graus dos vários PP em várias tarefas, de forma a determinar se os graus CH atribuídos por diferentes PP variam significativamente a cada TP executada, com 95% de confiança, e também determinar se cada diferente TP se traduz em graus CH diferentes com 95% de confiança.

Na pesquisa bibliográfica realizada, dentro das bibliotecas a que o DCTA tem acesso, não foram encontradas análises semelhantes sobre a validade da aplicação da Escala CH, em que pese seu vasto emprego na indústria e o longo tempo em utilização. O uso da técnica de Análise da Variância tampouco foi encontrado correlacionado com o emprego da Escala Cooper-Harper.

O trabalho está dividido em três seções, sendo esta primeira uma introdução ao assunto relacionado ao Ensaio em Voo e apresentação de metodologia. A segunda seção divide-se em quatro subseções, apresentando as fases de verificação prévia para a aplicação da Análise de Variância conforme [4]. A terceira seção apresenta as conclusões deste trabalho.

HANDLING QUALITIES RATING SCALE (COOPER-HARPER)

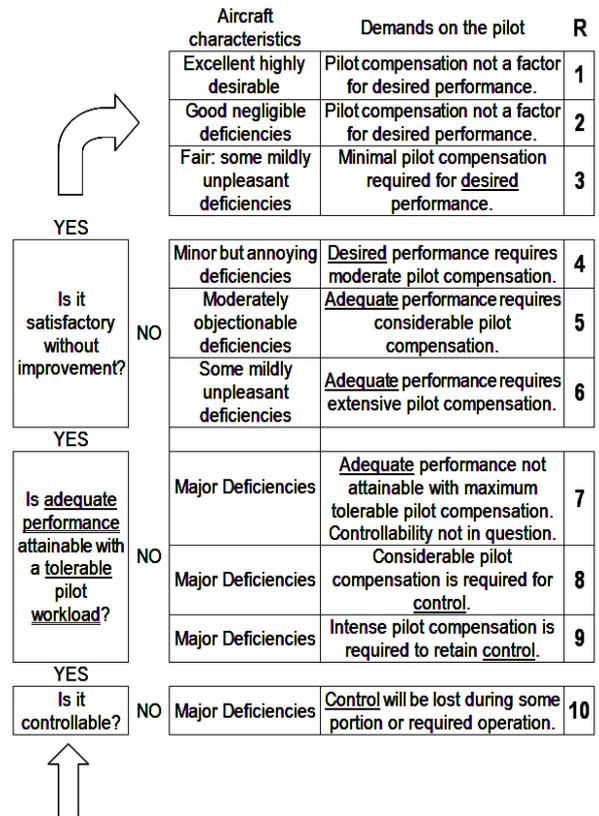


Fig. 1. Escala de Graus Cooper-Harper

II. ANÁLISE DOS DADOS COLETADOS

1) *Experimento*: Para o estudo da significância dos graus atribuídos por diferentes PP, foi desenhado um experimento que permitisse a verificação de graus atribuídos por seis PP nas mesmas TP.

Durante uma grande campanha de ensaios em voo, solicitou-se que, sempre que fosse seguro, o PP, coordenado pelo EP que gerenciava o voo de ensaio, deveria atribuir grau CH a cada tarefa de pilotagem elaborada, num total de treze tarefas. Ao final da campanha, foi possível elaborar uma tabela em que todos os seis PP atribuíram, pelo menos, dois graus CH a cada uma de sete tarefas de pilotagem. As demais seis TP não foram utilizadas por não contarem com um número mínimo de observações por parte dos PP ao longo da campanha.

A Tabela I mostra as tarefas de pilotagem desenvolvidas avaliadas com uma amostragem de dois graus CH.

TABELA I TAREFAS DE PILOTAGEM AVALIADAS

TPx	Descrição da Tarefa de Pilotagem
TP1	Aproximação de Pista deslocada com correção a 200 ft de altura
TP2	Decolagem
TP3	Manutenção de Reta durante o Táxi
TP4	Subida mantendo lei de velocidade
TP5	Pouso de Precisão
TP6	Execução de Curva durante o Táxi
TP7	Aproximação ILS

As TP foram desenvolvidas sob o critério de representarem tarefas de segmentos de voo bem diferentes, o que implicaria em resultados de controlabilidade e QDP distintos.

Os seis PP foram designados como PP1 a PP6, e as tarefas foram designadas como TP1 a TP7. A Tabela II mostra os graus atribuídos. Cada célula da Tabela traz os graus CH atribuídos pelos PP em cada voo da campanha.

TABELA II GRAUS CH ATRIBUÍDOS DURANTE O EXPERIMENTO

	PP1	PP2	PP3	PP4	PP5	PP6
TP1	3;3	3;2	7;2	5;5	5;5	5;5
TP2	3;3	2;3	5;2	3;3	3;2	3;2
TP3	3;3	4;2	2;2	2;2	2;5	2;2
TP4	3;3	5;2	5;2	4,5;4	5;5	3;3
TP5	3;3	3;2	2;2	3;2	5;3	4;4
TP6	4;3	4;3	4;4	4;4	5;5	3;3
TP7	4;4	3;3	5;4	5;5	2;1	2;2

No total, foram analisados 84 graus atribuídos por seis PP em sete tarefas de pilotagem, com duas repetições cada.

2) *Resultados iniciais:* Inicialmente, foi checada a dispersão dos graus atribuídos por todos os PP em todas as TP. A Fig. 2 mostra o gráfico *boxplot* da amostra de graus.

O gráfico da Fig. 2 mostra que, apesar de alguns pilotos apresentarem uma dispersão de resultados maior que outros, a distribuição desses resultados ocorre em graus semelhantes, sempre havendo intersecção entre os resultados de todos os PP. A semelhança nos *output* determina a necessidade de estudos mais detalhados, aprofundando a esta verificação inicia, a exemplo do que mostra [4].

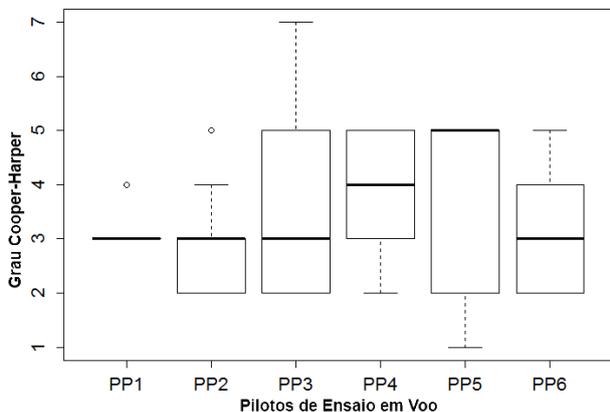


Fig. 2. Distribuição de graus atribuídos por cada PP

O *boxplot* da Fig. 3 mostra que os graus atribuídos a cada TP diferem razoavelmente uns dos outros, o que traz a suspeita de que cada tipo de Tarefa de Pilotagem resulta num grau significativamente diferente, o que faz sentido, pois cada tipo de tarefa apresenta, por hipótese, uma QDP diferente. O gráfico também aponta para a necessidade de um estudo mais detalhado.

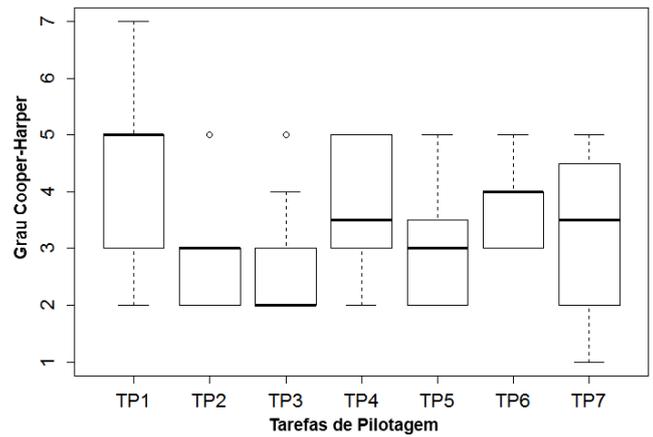


Fig. 3. Distribuição de graus atribuídos a cada TP

3) *Verificação da hipótese de normalidade dos resíduos:* uma vez estabelecida a necessidade de analisar mais a fundo os dados relativos aos graus CH, prosseguiu-se para a verificação da adequabilidade do modelo de Análise de Variância. Assim, foi verificada a hipótese de normalidade dos resíduos, conforme a Fig. 4, por meio do gráfico qqnorm.

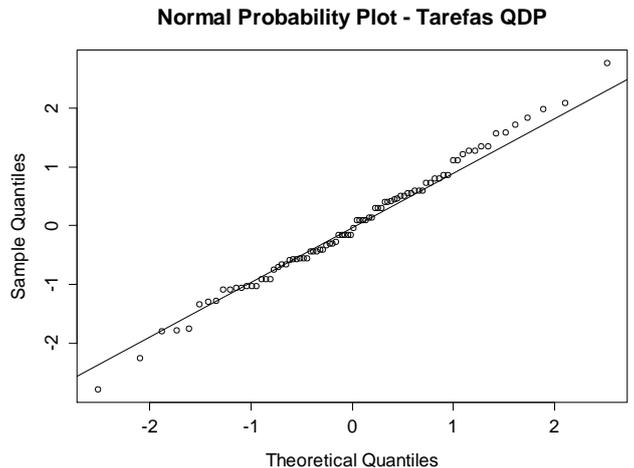


Fig. 4. Gráfico de verificação de normalidade de resíduos

O gráfico da Figura 4 mostra que a distribuição dos resíduos dos graus CH se aproxima da distribuição normal, indicando a viabilidade do uso do modelo para a Análise de Variância (ANOVA).

Prosseguindo para outra forma de verificação, foi realizada uma verificação objetiva da normalidade da distribuição dos resíduos pelo teste de normalidade de Shapiro-Wilk. O *output* obtido é o copiado na Figura 5.

```
> shapiro.test(anova$residuals)
Shapiro-wilk normality test
data: anova$residuals
W = 0.9943, p-value = 0.9749
```

Fig. 5. Resultado do teste de normalidade de Shapiro-Wilk

De acordo com [5], o resultado do teste indica uma grande similaridade com a distribuição normal, com um valor de W muito próximo a 1, além de um p -value bastante acima de 5%, que indica que a hipótese nula do teste de Shapiro-Wilk (de que a distribuição é normal) não pode ser rejeitada. Ainda que esta não seja uma confirmação de que a distribuição é efetivamente próxima à normal, os grandes valores de W e p , junto ao $qqnorm$ da Fig. 4, dão confiança para prosseguir para a ANOVA da distribuição.

4) *Análise de Variância e determinação de significância de parâmetros:* Após as verificações iniciais, foi levantada a Análise de Variância dos graus atribuídos frente aos dois parâmetros de estudo, o PP responsável pelo grau e a TP em avaliação, com o resultado conforme o descrito na Fig. 6.

```
> anova = aov(nCH~fPP+fTsk)
> summary(anova)
          Df Sum Sq Mean Sq F value Pr(>F)
fPP         5   8.09   1.617   1.327 0.2624
fTsk         6  24.08   4.013   3.294 0.0064 **
Residuals   72  87.73   1.218

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
                '.' 0.1 ' ' 1
```

Fig. 6. Resultados da Análise de Variância dos graus atribuídos

A Análise de Variância indica que a significância do fator Piloto de Ensaio em Voo tem uma confiabilidade menor que 74% ($Pr(>F)=0.2624$), abaixo do critério de 95% de confiança definido inicialmente. Assim, não se pode afirmar, com 95% de confiabilidade, que diferentes Pilotos de Ensaio em Voo atribuem graus diferentes, implicando no uso de critérios diferentes. Este resultado implica que o fator PP tem baixa significância no resultado grau CH, para as TP investigadas. O resultado corrobora, de certa maneira, o previsto por G. E. Cooper e R. P. Harper em 1969 [2].

Já o fator Tarefa de Pilotagem tem um grau de significância de mais de 99% de confiança. Dessa forma, pode-se afirmar que diferentes Tarefas de Pilotagem recebem graus CH diferentes, o que demonstra que efetivamente um piloto bem qualificado retorna uma avaliação criteriosa de cada Tarefa de Pilotagem a mais de 99% de confiança, acima do critério estabelecido inicialmente. Este tipo de resultado já era esperado, conforme citado por G. E. Cooper e R. P. Harper [6], pois as qualidades de pilotagem de uma aeronave são dependentes da tarefa de pilotagem escolhida.

III. CONCLUSÃO

A análise do Experimento desenhado com base nos resultados da campanha permite reforçar a importância da utilização do método de escalas padronizadas para a avaliação de QDP. A qualificação de Pilotos de Ensaio em Voo na determinação das características de um sistema, padronizando e uniformizando os graus e os comentários também pode ser apontado como uma característica importante.

Das análises realizadas, conclui-se que, para um nível de confiança de 95%, o grau CH atribuído por um PP não difere significativamente dos graus atribuídos pelo um grupo de seis

PP. Dessa forma, em campanhas de ensaio em voo com um pequeno número de voos, exemplo de campanhas de integração de armamentos ou de certificação de variantes de um modelo de aeronave, o EP gerente da campanha pode planejar a mesma considerando a possibilidade de que um ou dois PP poderão retornar um resultado de avaliação das QDP da aeronave ensaiada num grau de confiança equivalente ao de um grupo maior de PP. Dentro do cenário atual e constante de escassez de EP e PP, isto permitiria a execução de algumas campanhas com o emprego de um menor número de PP.

A contrapartida, nesse caso, é que o grau CH é somente parte da avaliação, devendo ser complementado pelas opiniões subjetivas do piloto. Assim, um pequeno número de PP gerariam um pequeno número de opiniões de pilotagem, o que obrigaria um maior trabalho de análise e padronização por parte dos EP encarregados dos voos de ensaio. Este é um tipo de problema bastante comum em campanhas de modificações de aeronaves, onde o número de voos é geralmente pequeno e a avaliação das QDP da aeronave pode, efetivamente, repousar num número bastante reduzido de pilotos ou de repetições de tarefas, com chances de não ser possível coletar resultados em mais de um voo.

Mesmo durante o experimento realizado para a elaboração deste Artigo, foram desenvolvidas treze tarefas de pilotagem inicialmente, porém, somente foi possível coletar dados suficientes para a análise comparativa de sete dessas tarefas, ao longo de quase 100 vôos.

Por fim, apesar de ter sido possível determinar o resultado estatístico com elevada confiança, é interessante que o experimento fosse continuado junto ao IPEV (Instituto de Pesquisas e Ensaio em Voo), coletando dados de uma maior quantidade de equipagens de ensaio em voo em Tarefas de Pilotagem de forma confirmar o resultado com outros grupos de PP e elevar a confiança do uso de menores amostragens de graus CH.

REFERÊNCIAS

- [1] G.E. Cooper e R.P. Harper (1966) "A revised pilot rating scale for the evaluation of handling qualities". AGARD Flight Mechanics Panel. Cambridge, England.
- [2] G.E. Cooper e R.P. Harper (1969) "The use of pilot rating in the evaluation of aircraft handling qualities". NASA Technical Note TN D-5153. NASA, EUA.
- [3] Brasil (2004) "Apostila E-QA06 TEV Qualidades de pilotagem". IPEV.
- [4] D. C. Montgomery (2001) "Design and analysis of experiments", 5th edition. John Wiley & Sons, USA.
- [5] P. Dalgaard (2006) "[R] shapiro.test() output" Disponível em <<https://stat.ethz.ch/pipermail/r-help/2006-July/109123.html>>. Acesso em 29/06/2015.
- [6] R.P. Harper e G.E. Cooper (1984) Handling Qualities and Pilot Evaluation. 1984 Wright Brothers Lectureship in Aeronautics. EUA.