

# 宇宙航空研究開発機構研究開発報告

## JAXA Research and Development Report

---

### Threat and Error Managementを考慮した CRMスキル計測手法の開発

飯島 朋子, 津田 宏果, 野田 文夫

2011年9月

宇宙航空研究開発機構

Japan Aerospace Exploration Agency

## 目 次

1.	はじめに .....	2
2.	提案のCRMスキル計測手法 .....	3
2.1	はじめに .....	3
2.2	CRMスキル計測 .....	3
2.3	CRMスキル計測の前提条件 .....	3
2.4	データの秘匿性 .....	4
	(1) 計測者・被計測者 .....	4
	(2) 計測手法を各会社を導入する場合 .....	4
2.5	CRMスキル計測シート .....	4
	(1) 計測項目前提条件 .....	4
	(2) 評点尺度前提条件 .....	4
	(3) 計測手順 .....	4
	(4) 計測シート記入方法 .....	5
	(5) 評点の際の注意点 .....	5
2.6	計測者の養成 .....	6
	(1) 計測者の資格要件 .....	6
	(2) 計測者の標準化 .....	6
2.7	Threat and Error Managementを考慮したLOFTシナリオ <sup>9)</sup> .....	6
3.	CRMスキル計測実験 .....	9
3.1	概要 .....	9
3.2	計測者 .....	9
3.3	模擬LOFTシナリオ .....	9
3.4	実験手順 .....	9
4.	結果と分析 .....	9
4.1	全体傾向 .....	9
4.2	提案の計測シートの有効性 .....	12
4.3	計測者同士の議論の有効性 .....	12
4.4	計測者の評点傾向一致の度合い .....	15
4.5	各計測者の評点傾向 .....	18
4.6	計測者間のばらつきが大きいCRMスキル計測項目 .....	20
4.7	空欄の目立つCRMスキル計測項目 .....	22
4.8	当該機種 of 操縦経験有無による相違 .....	24
4.9	LOFTシナリオの有効性 .....	24
	(1) Threatの数との関連 .....	24
	(2) Threatの種類との関連 .....	24
	(3) シナリオ評価 .....	24
4.10	チームの総合点 .....	25
5.	考察 .....	26
5.1	CRMスキル計測シート .....	26
	(1) Threat別vs. 飛行フェーズ別計測シート .....	26
	(2) CRMスキル計測項目自体の解釈 .....	27
	(3) 空欄の目立つCRMスキル計測項目 .....	27
	(4) チームの総合点と各CRMスキル項目の評点との関連 .....	27

5.2	計測者の標準化.....	27
(1)	True Score導出のための計測者同士の議論.....	27
(2)	計測者の資格要件 .....	28
(3)	今後の計測者標準化に向けて .....	28
5.3	LOFTシナリオ開発.....	29
6.	まとめ .....	29
7.	参考文献 .....	30
8.	謝辞 .....	30
9.	付録 .....	30
	付録 1 CRMスキル計測シート (Threat別) .....	31
	付録 2 CRMスキル計測シート(飛行フェーズ別).....	53
	付録 3 JAXA CRMスキル観測シート (例) .....	63
	付録 4 ANOVA結果.....	65

## 略語

AIL	Aileron
ANOVA	ANalysis Of VAriance
A/P	Auto Pilot
ATEC	Association of Air Transport Engineering and Research
A/T	Auto Throttle
ATC	Air Traffic Controller
BRFING	Briefing
CAP	Captain
CB	Cumulonimbus
CONT	Control
CRM	Crew Resource Management
CRZ	Cruise
DAL	Descend / Approach / Land
FMS	Flight Management System
FO	First Officer
FWD	Forward
ICAO	International Civil Aviation Organization
JAM	Jamming
JAXA	Japan Aerospace Exploration Agency
LOFT	Line Oriented Flight Training
LOSA	Line Operations Safety Audit
LSD	Least Significant Difference
MOD	Moderate
PF	Pilot Flying
PNF	Pilot Not Flying
RA	Resolution Advisory
SEV	Severe
SOP	Standard Operating Procedure
TA	Traffic Advisory
TCAS	Traffic Alert and Collision Avoidance System
TEM	Threat and Error Management
TOC	Take off / Climb
TURB	Turbulence
WX	Weather

# Threat and Error Management を考慮した CRM スキル計測手法の開発\*

飯島 朋子<sup>\*1</sup>, 津田 宏果<sup>\*1</sup>, 野田 文夫<sup>\*1</sup>

## Development of CRM Skills Measurement Method including Threat and Error Management Concept\*

Tomoko IJIMA<sup>\*1</sup>, Hiroka TSUDA<sup>\*1</sup> and Fumio NODA<sup>\*1</sup>

### ABSTRACT

The Japan Aerospace Exploration Agency (JAXA) has developed a Crew Resource Management (CRM) Skills measurement method that includes a Threat and Error Management (TEM) concept and identifies a crew's level of CRM Skills by the way in which they manage human errors and threats. To validate the proposed method, a CRM Skills measurement experiment was carried out by four raters observing pre-recorded simulated LOFT (Line Oriented Flight Training) scenarios carried out in cooperation with airline companies. The experiment included trial discussions between the raters to introduce a “True Score<sup>5)</sup>” which contributes to “inter-rater-reliability”. The results of the experiment showed that the proposed CRM Skills measurement sheet assisted raters in evaluating crew behavior when managing or mismanaging threat / errors from a consistent viewpoint, but individual differences of assessment between the raters still remained. Additionally, the “True Score” discussion prompted raters to clarify their evaluation rationales. As a whole, it was found that when the average values of CRM Skills items are lower, the differences between rater evaluations are greater. It is possible that factors contributing to these individual differences include not only the contents of CRM Skills measurement sheet and the inter-rater-reliability training method (the “True Score” discussion process), but also types of threat included in the simulated LOFT scenarios and the timing of their appearance.

Keywords : Crew Resource Management, Behavioral Markers, CRM Skills Measurement, Line Oriented Flight Training, Threat and Error Management, Inter-Rater-Reliability

### 概 要

Threat and Error Management (TEM) を考慮した Crew Resource Management (CRM) 訓練の効果を測定するための CRM スキル計測手法を開発した。本手法の妥当性を検証するため、航空会社の協力を得て模擬 Line Oriented Flight Training (LOFT) を行い、撮影した模擬 LOFT のビデオを使って、選出された 4 名の計測者による計測実験を実施した。計測実験においては、試験的に「True Score<sup>5)</sup>」を出すための計測者同士の議論の場を取り入れた。実験の結果、提案の CRM スキル計測シートは、計測者のクルー行動に対する着眼点の一致を促したが、依然として、計測者間のばらつきは存在した。また、計測者間で議論を行なうことにより、計測者の評点の考え方に対する観点をそろえることができた。全体的に計測項目の評点の平均値が低い場合は計測者間のばらつきが大きくなる傾向であることが分かった。ばらつきの要因として、CRM スキル計測シートの作成方法、計測者の標準化訓練方法（「True Score<sup>5)</sup>」導出のための計測者同士の議論の場）の他に、LOFT シナリオに組み込む Threat の種類と、それを入れるタイミングが関与している可能性があることが明らかとなった。

---

\* 平成 23 年 7 月 11 日 受付 (Received 11 July 2011)

\*1 航空プログラムグループ 運航・安全技術チーム

(Operation and Safety Technology Team, Aviation Program Group)

## 1. はじめに

航空機事故原因の 5 割以上がヒューマンエラーによるものとされており、事故の主要因は Crew Resource Management (CRM) スキルの不足であると言われている。ここで CRM とは、利用可能な資源（ハードウェア、人間、情報）を最大限に活用して安全な運航を実現することであり、CRM スキルとは CRM を実践する能力と定義される。言い方を変えれば、コミュニケーションやチームワークを向上させて、ヒューマンエラーを防ぎ、たとえヒューマンエラーが生じてもクリティカルな状況や事故に至らないように、チームで適切に業務を遂行していくことである。CRM を強化することが航空機事故防止に有効であることから、現在我が国においても航空機乗務員に対する CRM 訓練が義務化されている。

CRM 訓練の効果を高めるためには、CRM セミナーや Line Oriented Flight Training (LOFT) において CRM の重要性や取り組み姿勢を理解させるだけでなく、ライン運航や教育訓練の場で CRM スキルを定着化させる必要がある。CRM スキルをより実践的なものとして定着化させるために、CRM を実践するための乗員の行動指標（指標として明確に示される具体的な行動）を新たに作成し、訓練プログラムに導入する必要がある。そこで宇宙航空研究開発機構（以下 JAXA）では大手航空会社の協力の下に我が国の特徴に合致した行動指標を開発し提案した<sup>1)</sup>。

しかしながら、安全性と経済性への要求が一層厳しくなっている今日、さらな

る訓練の高度化と効率化のためには、行動指標を取り入れた訓練プログラムの導入のみならず、その訓練によって何が身に付き、何が足りないかを計測する手法が必須である。これにより CRM 訓練の効果が計測可能となれば、問題点や改良点を訓練プログラムにフィードバックすることができる。そこで JAXA では、CRM 訓練の効果を計測するための CRM スキル計測手法を開発し、改良を重ねてきた<sup>2)</sup>、<sup>3)</sup>。文献(3)では、飛行フェーズごとに CRM スキル計測を実施したところ、9 名の計測者においてばらつきがみられた。その主な原因は着目する行動が異なるというものであった。評点結果にばらつきが生じれば、データの信頼性を損ない CRM 訓練プログラムへの適切なフィードバックが出来なくなるため、計測者間での観点や判断基準の統一すなわち標準化訓練の必要性が指摘された。飛行フェーズごとに CRM スキルを計測する他の手法に Line Operations Safety Audit (LOSA)<sup>4)</sup>があるが、ここでも計測者の訓練が実施されている。

また、今日ではヒューマンエラーは避けられないという考えの下で Threat and Error Management (TEM) の概念を運航乗務員の資格・訓練要件に追加する国際的な取り組みがすすめられ、ICAO 第 1 及び第 6 付属書の改正が行われている（2006 年 11 月適用）。ICAO は第 6 付属書において、運航乗務員の訓練プログラムとして TEM を含む人間の能力と限界に関する訓練を定期訓練でも実施すること、及びその能力（Competency）の評価

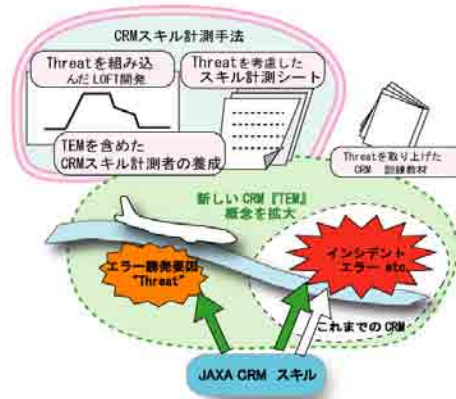


図 1 TEM を考慮した CRM スキル計測手法概念図

(Assessment) を含むことを国際標準で定めている。このことから、CRM スキル発揮の目的が Error Management だけではなく、Threat (エラーを誘発する要因、運航阻害要因) Management も含めることに変わりつつあり、今後 CRM 訓練の効果を測定する場合においても、TEM を考慮した CRM スキルを計測する手法が必要になると考えられる (図 1 TEM を考慮した CRM スキル計測手法概念図参照)。

そこで本研究では、計測者の着目する行動を統一することで計測者の評点のばらつきを少なくし、かつ TEM の概念を取り入れた計測を実施するために、シナリオに設定された Threat ごとに CRM スキルを計測する手法を提案する。また計測者標準化訓練の有効性を検証するため、試験的に計測者の「True Score<sup>5)</sup>」を導出する実験を試みた。True score とはビデオテープに録画されたクルーの行動について、専門家達がそれぞれ評点付けを行い、その後、議論を通して全員が合意して出された数値である。

本稿では JAXA が提案する CRM スキル計測手法及び手法の有効性を検証するために行った実験について述べる。本実験で検証した計測手法は今後、我が国の全ての航空会社が TEM を含めた有効な CRM 訓練を実施していくための利用可能なツールとして幅広く普及をめざすものである。

なお本研究は、国土交通省航空局技術部運航課の委託を受けて、航空輸送技術研究センター (ATEC) を事務局とした「ヒューマンエラーを考慮した航空従事者等教育訓練方式の確立に関する調査・研究委員会」<sup>6)</sup>の研究の一環 (TEM の概念を取り入れた CRM 訓練の実施と運航乗務員の CRM スキル定着度を計測する手法の調査・研究) として実施したものである。

## 2. 提案の CRM スキル計測手法

### 2.1 はじめに

CRM スキル計測手法開発の際には、「CRM スキル計測シート」、信頼するデータを得るための「計測者の養成」、TEM のための CRM スキルを発揮する有効なシナリオである「LOFT シナリ

オ」を検討する必要がある。以下に本研究で提案する CRM スキル計測方法、計測の前提条件と上記要素について述べる。

### 2.2 CRM スキル計測

本研究で定義する CRM スキル計測とは、計測者がある場面の (例えば、離陸時の) ある行動を見て (例えば、Tire Burst 時)、観測した CRM スキルに、CRM スキルの実践度の点数をつけることである。ここで、着目する行動が異なったり、点数が異なったり等の計測者間でばらつきが生じることが問題となっている。<sup>3)</sup>ばらつきの原因は、着目する行動、CRM スキル、実践度が計測者間で異なるということである。例えば計測者 A は、Take off 時のタービュランスに対して対処した行動を見る、また計測者 B は、異なる場面の異なる行動を観測するといったように着目する行動が異なったり、また、着目する場面、行動が同じであったとしても、着目する CRM スキルが異なっていたり、ある計測者は、低い評点としたり、ある計測者は高い評点とするといった問題がある。

そこで、計測者間のばらつきを少なくするために、計測者には、Threat や Error に対処した際のクルーの行動のみが評価可能となるように、CRM スキル計測シートを改良した。これにより、TEM の概念を考慮し、計測者の観測する行動を一致させることができると考えられる。

また、CRM スキルの観測や、評点のばらつきを抑えるために、計測者同士の認識を一致させるための議論を実施した。これは、着目している CRM スキルの違いや、評点の違いを議論して、最終的には、このスキルに着目して何点とするといった一連の作業 (「True Score<sup>5)</sup>」導出の試み) を意味する。

### 2.3 CRM スキル計測の前提条件

本手法は、チェックあるいは審査といった乗員個人のスキル計測を対象とせず、ある組織全体を対象とした CRM スキル定着化分析を行い、問題点や改良点を次の訓練プログラムにフィードバックするこ

とを目的とする。

本目的の上で満たすべき前提条件を以下に述べる。

## 2.4 データの秘匿性

本計測は、個人の評価を対象とするのではなく全体的な傾向を分析するために行うので、抽出データは匿名扱いとする。

### (1) 計測者・被計測者

- ・ 被計測者はセットクルーとする。
- ・ LOFT や日常運航で、2.5 に述べる資格要件を満たす計測者が計測する。

### (2) 計測手法を各会社に導入する場合

文献(6)において、計測のための技術的課題だけではなく、組織的な課題が指摘された。本計測手法を各会社に導入する場合には、以下を満たすことを提案する。

- ・ CRM スキル計測は、乗員個人の CRM スキルを評価するために導入するのではなく、CRM 訓練システムにおける品質管理の構成要素として導入すべきである。
- ・ 乗員が CRM スキル計測について正しい知識を持ち理解していること、本手法導入に対して集团的合意が達成されている必要がある。
- ・ 計測者・被計測者と計測を運用する組織内の信頼形成を満たす必要がある。

## 2.5 CRM スキル計測シート

提案の CRM スキル計測シートを付録 1, 2 に示す。本手法は、2.6 に述べる選ばれた計測者が、セットクルーの CRM スキルレベルを付録 1, 2 に示す「計測指標（以後、計測項目）」ごとに 4 段階で評点付けし、コメントを記入する主観評価を基本としている。付録 1, 2 はそれぞれ Threat 別、飛行フェーズ別に CRM スキルを計測する主観評価シートを示している。

Threat 別の CRM スキル計測シートは、計測者の着目する行動を「Threat に対処した際の行動」に統一することで計測者の評点のばらつきを少なくし、かつ TEM の概念を取り入れた計測を実施するために提案した。飛行フェーズ別シートは Threat 別と比較するために用意した。

### (1) 計測項目前提条件

- ・ クルーの行動が観測可能なものであること

計測項目は過去に JAXA が提案した計測指標<sup>2), 3)</sup>を基に、明確に行動として表れ、計測者が目や耳で観測できる行動のみを選択する。（例えば「ストレスをコントロールする」は観測出来ないので除外する）。

- ・ 評価の境界が明確であること  
文献(3)で指摘された観測が難しい行動及び行動の分類が困難なスキルについては、計測項目として一緒にした（警戒・予測、プランニング・優先順位付け）。ただし行動指標との対応の便宜を考えて、シート上の「計測ポイント」に行動指標に関する文章は残した。
- ・ 計測項目の説明（付録 1, 2 に示す計測ポイント）  
計測項目の説明は計測者全員が理解できること。計測者ごとに考え方がばらつかないように定義や計測のポイントなどを明確化する必要がある。
- ・ 実運用性と解析のし易さの考慮
  - 計測者に出来るだけ負担にならず、かつ解析しやすい項目とする。
  - 計測項目が多いと計測者の負担が大きくなるため、シンプルな計測項目となるよう考慮する。

### (2) 評点尺度前提条件

CRM スキルを評点する際、1 点から何点までの何段階評価が相応しいのか、正解は決まっていない。ただし、3 段階や 5 段階など中央の値が存在する評点段階では、計測者の結果が中央値に偏る危険性がある。

そこで、本研究では 4 段階の評点を採用した。評点尺度の定義は、カンタス航空の LOFT 後のデブリーフィングで用いられているシート<sup>7)</sup>を参考に作成した。表 2-1 に評点尺度のガイドラインを示す。これは、評点付けする際のガイドラインとして計測者に提示するために用意した。

### (3) 計測手順

- 模擬 LOFT におけるクルーの行動をビデオから観察し、CRM スキル観測シート<sup>2), 3)</sup>（付録 3）に抽出した行動についてメモを取る。

- b. ビデオ終了後、観測シートを見ながら CRM スキル計測シートに評点・コメント・全体コメント・チーム総合点を記入する。

#### (4) 計測シート記入方法

##### a. Threat 別計測シート（付録 1）

##### ① 設定された Threat に対する CRM スキル実践度評点

本シートでは、フライト中に遭遇・発生する Threat について、(2)に述べた 4 段階で 14 項目（計測項目）への評点付けを行う。例えば、Threat が 6 個発生するシナリオの場合、計測者は 14 項目×6 個＝84 箇所への記入を行う。Threat 別の計測であることから、各シートの冒頭にその Threat が発生する飛行フェーズと、Threat の内容を予め記載しておく。このため、Threat 別計測シートは、各シナリオによって枚数が異なる。付録 1 に示す Threat の括弧書きの内容説明は、Error や Threat の分類コードである LOSA Error Code Book 9.0<sup>8)</sup>を参照したものである。

##### ② 上記評点に対するコメントを記入

評点の理由について以下の観点で具体的に記述する（評点 2 以下、4 については要記入）。

- クルーは Threat や Error が生じた際、
- －どのように行動したか？
  - －その行動は適切・不適切だったか？
  - －行動は効果につながったか（パフォーマンスに現れたか）？
  - －行動自体が行われていなかったか？

##### ③ 全体コメント・チーム総合点を記入

フライト全体（Overall）においては、全体コメント、機長、副操縦士へのコメント、その他・Standard Operating Procedure (SOP)遵守に関するコメントを記入する。最終的にはチームの総合点を 4 段階で記入する。

##### ④ 計測者が自ら抽出した Threat / Error の取り扱い

計測者が抽出した Threat や Error があれば別のシートに抽出した Threat と飛行フェーズ、それに対する評点、及びコメントを記入する。

##### b. 飛行フェーズ別計測シート（付録 2）

飛行フェーズ別計測シートでは、フライトを 4 飛行フェーズ（Pre departure / Taxi Out, Take Off / Climb, Cruise, Descend / Approach / Land）に分け、それぞれの飛行フェーズごとに計測項目の評点付けを行う。計測項目は、Threat 別計測シートと同じ、全 14 項目であり、各飛行フェーズにおいて評点に関するコメントを記入する。フライト全体においても上記に述べた Threat 別計測シートと同様である。

#### (5) 評点の際の注意点

a. 観測された行動に関して評点する。観測できなかった行動については空欄にする。ただし、被計測者が「何も行動しなかった」と「観測できなかった」とは異なることに注意する。すなわち TEM を実施するために取るべき行動があるのに、その行動が行なわれず、観測されなかった場合には(2)に述べた評点尺度に従い、「エラー発生に寄与する可能性有り」で 1 点となる（例：安全への主張をすべき所で、何も言わなかった場合等）。

b. 評点に迷った場合（例：2 点か 3 点か？）

- CRM スキル行動指標<sup>1)</sup>を参照する。CRM スキル行動指標は、CRM スキル観測シート（付録 3）の裏に提示されている。
- Error が観測されても、チームで気付いて対処出来れば 3 点になる計測項目もある。
- CRM スキルの評点であって、テクニカルスキル（操縦技術）の評点ではないことに注意する。

##### c. 評点で犯しやすいエラー

文献(5)で述べられている評点で犯しやすいエラー（Halo Error, Leniency Error, Severity Error, Central Tendency Error）に加えて、これまでの研究<sup>2), 3), 6)</sup>から見られた以下の傾向を注意点としてあげる。

- 合否の判定との勘違い

Leniency（寛大さ）Error に類似するものであるが、CRM スキルは悪かったのに



2点、1点への評点を躊躇する場合がある。CRM 訓練効果の評価と改良すべき点の抽出に主眼をおいていること、及び2点、1点は落第点ではないことを理解して正確に評点する必要がある。

- 第一印象に左右される

Halo（後光） Error に類似するが、Flight Preparationなどで被計測者の CRM スキル実践度が悪いと、その第一印象に影響されて後々の計測結果も悪くなる可能性がある。逆に第一印象が良ければその後の計測結果が良くなる傾向がある。

## 2.6 計測者の養成

### (1) 計測者の資格要件

文献(3)、(6)の結果から、計測者の資格要件を以下に定めた。

- ・ CRM 関連部署の経験を有している。
- ・ LOFT 当該機種種の Standard Operating Procedure (SOP)に関する知識を有している。
- ・ LOFT 当該機種種の操縦経験が豊富であることが望ましい。
- ・ 計測に使用する LOFT シナリオ、シナリオに含まれる Threat を予め知っている。
- ・ 計測シートの使用方法を慣熟している。
- ・ 用いる行動指標を理解している。
- ・ 標準化訓練を受けている。
- ・ その組織における、CRM に対する“Company Policy”を理解している。

### (2) 計測者の標準化

- ・ 本研究で定義する CRM スキル計測とは、前述したように計測者がある場面の（例えば、離陸時）ある行動を見て（例えば、Tire Burst 時）、観測した CRM スキルに、CRM スキルの実践度の点数をつけることである。ここで、着目する行動が異なったり、点数が異なったり等の計測者間でばらつきが生じることが問題となっている<sup>3)</sup>。このばらつきを小さくして全計測者間で標準化がなされていることが望ましい。標準化がなされていなければ評点結果にばらつきが生じ、データの信頼性を損ない、CRM 訓練プログラムへの適切なフィードバックが困難となる。

そこで計測者の標準化訓練をまず先に実施する必要がある。標準化とは先にも述べたように、被計測者のチーム行動を観察し、実践されるべき CRM スキルに着目し、チームの CRM スキル実践レベルを評点に換算するといった一連の作業について、計測者同士の観点や認識を一致させることである。

- ・ 計測シートを用いた LOFT 計測者用ガイド、標準化するためのガイド・教育が必要である。本研究では、これまでの研究から評点付けのガイド・教育には以下を含むものとして提案する。例えば、以下の項目については、それぞれ正解はないが、重要なのは計測者間で共通に以下の項目やルールを理解していることである。理解が共通でなければ、評点にも影響が及び、点数がばらつく一因となる。
  - 評点の際に犯しやすいエラーの知識
  - 可否の判定ではないことの認識
  - 計測シートの基となる行動指標<sup>1)</sup>の教育
  - 各評点の持つ意味の理解
  - 計測するタイミングの統一（Threat 発生時／Threat 対処時）
  - エラーを犯した場合の判断（エラーを犯した以上は減点／気付いて対処出来れば減点せず）
  - 計測する対象の理解（個人としての乗員／チームとしての乗員）
  - SOP を計測対象とする／しない

## 2.7 Threat and Error Management を考慮した LOFT シナリオ<sup>9)</sup>

国内大手航空会社ではすでに Threat を取り込んだシナリオを作成して LOFT を実施していることから、CRM スキル計測に使用するシナリオは既存のシナリオから選択した。選択の観点として、Threat はシステムの故障だけに偏らずできるだけ多くの種類を含むものとしたが、Threat 挿入のタイミングに関して若干の相違がある。また、場合によっては教官が適宜 Threat を追加できるものとした。

その結果、A 会社が 2 つ、B 会社が 2 つの合計 4 つのシナリオが選択され、4 回の模擬 LOFT を実施した。シミュレータ

表 2-1 評点尺度ガイドライン

1 点	エラー発生に寄与する行動が観測される．すぐにこの行動を改善する必要がある．
2 点	改善すべき CRM スキルが観測される．または，CRM スキルが発揮されたがタスクの実践を助けていない．改善が必要である．
3 点	クルーメンバがライン運航で達成すべき CRM スキルを発揮．その行動は深刻なエラーに至らないように，効率的にタスクの実践を助ける．
4 点	全てのクルーメンバがライン運航で努力して達成すべき CRM スキルを発揮．その行動は通常必要とされるよりも効率的にタスクの実践を助ける．

注：3 点が標準

表 2-2 シナリオ構成

飛行フェーズ シナリオ		P1	P2	P3	P4	
	Threat の種類	件数				合計
シナリオ 1/ 機種 1	Aircraft	0	0	1	2	3
	Departure	0	0	0	0	0
	Arrival	0	0	1	1	2
	Cabin	0	1	0	0	1
	Operational	0	0	0	0	0
	ATC	0	0	0	0	0
	合計	0	1	2	3	6
シナリオ 2/ 機種 2	Aircraft	0	2	1	0	3
	Departure	0	1	0	0	1
	Arrival	0	0	0	0	0
	Cabin	0	0	1	0	1
	Operational	0	0	0	1	1
	ATC	0	0	0	0	0
	合計	0	3	2	1	6
シナリオ 3/ 機種 3	Aircraft	0	1	0	0	1
	Departure	0	1	0	0	1
	Arrival	0	0	1	0	1
	Cabin	0	0	1	0	1
	Operational	0	0	0	1	1
	ATC	1	0	0	0	1
	合計	1	2	2	1	6

注 P1: Pre departure / Taxi out, P2: Take off / Climb, P3: Cruise, P4: Descent / Approach / Land

はグラスコックピットである 3 機種を選択し、のべ 12 名の乗員が参加した。4 つのシナリオの内、CRM スキル計測実験に使用した 3 つのシナリオにあらかじめ設定された Threat の種類を表 2-2 に、シナリオの概観を図 2-1 から 2-3 に示す。

表 2-2 に示されるように、シナリオ 3 は他シナリオに比べて様々な種類の Threat が満遍なく入っていた。シナリオ 1、

2 においては Aircraft Threat が 3 つ入っているが、シナリオ 2 の場合は Aircraft Threat が Take Off 飛行フェーズで立て続けに入っており、シナリオ 1 の場合は Aircraft Threat が Descent / Approach / Land 飛行フェーズにおいて続くものの、本飛行フェーズにおいて Aircraft Threat の間に Arrival Threat が挿入されるという構成であった。

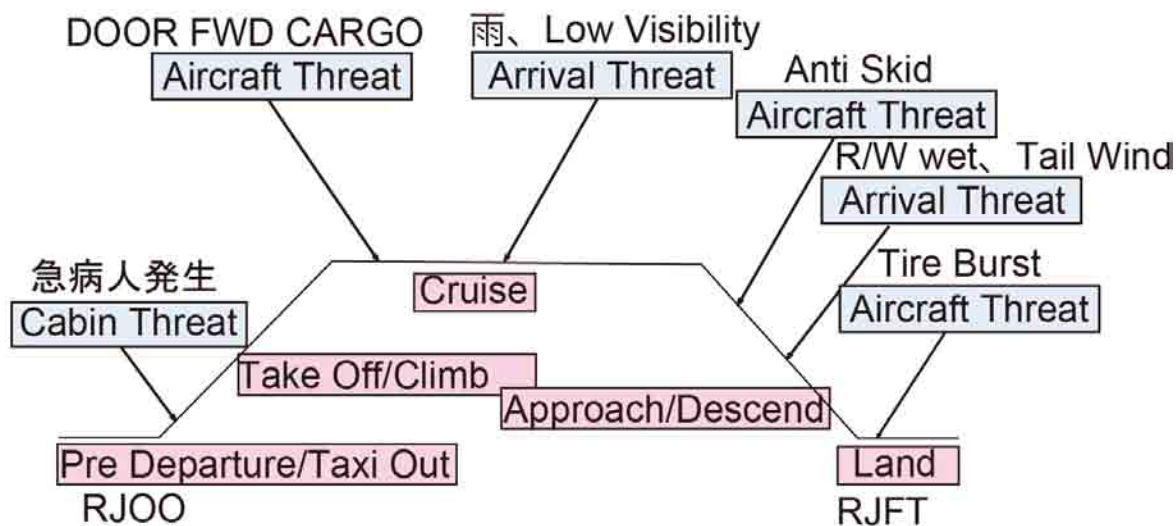


図 2-1 シナリオ 1

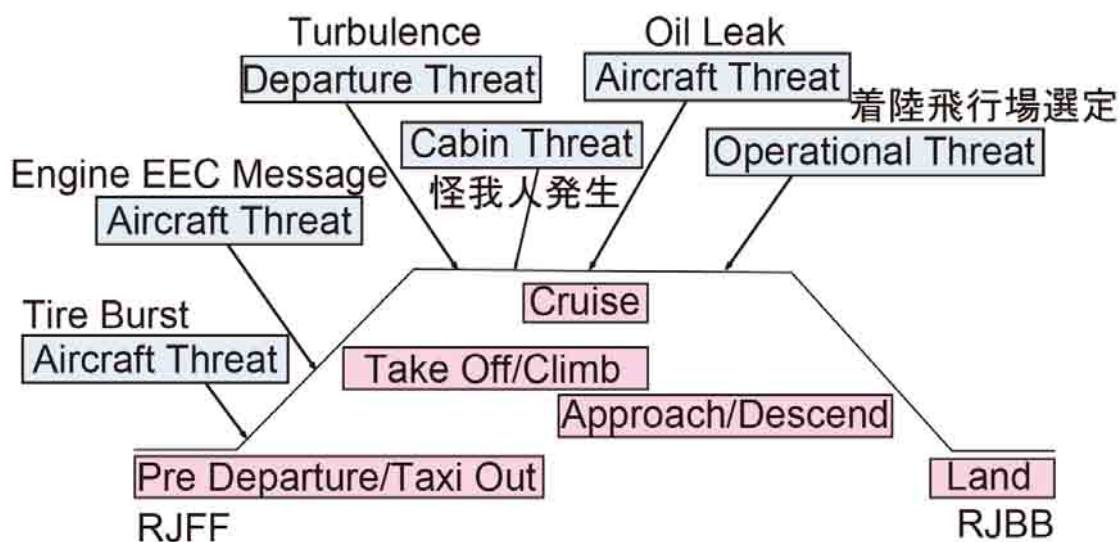


図 2-2 シナリオ 2

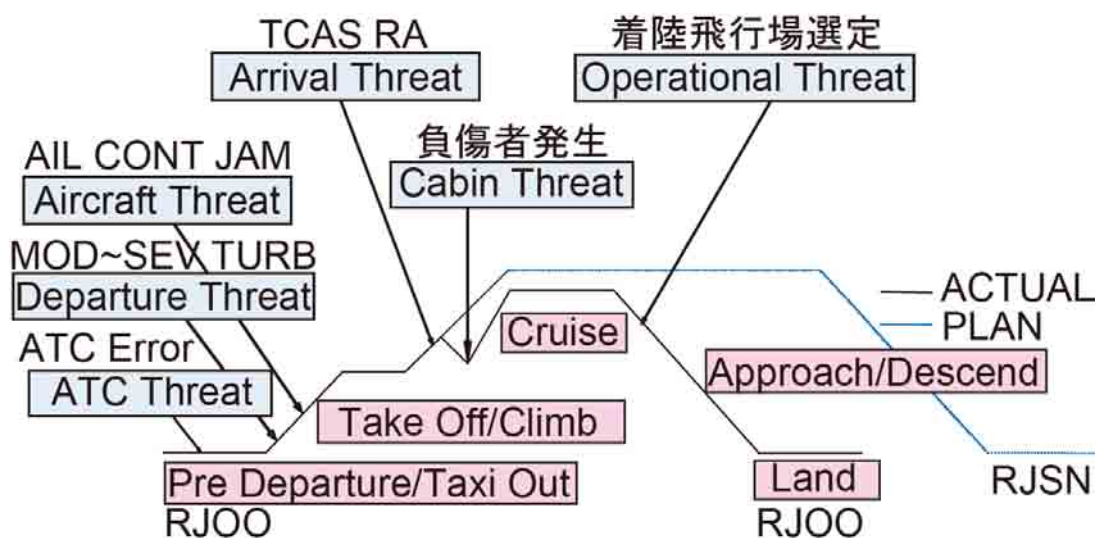


図 2-3 シナリオ 3

### 3. CRM スキル計測実験

#### 3.1 概要

提案の計測シート、計測者の「True Score<sup>5)</sup>」導出試験の有効性を検証するために、計測者による CRM スキル計測実験を実施した。実験の前に国内の航空会社 2 社において模擬 LOFT を実施し、2.7 に示す 3 つのシナリオのビデオを取得した。計測実験では、そのビデオを用いて、各航空会社で選出された 4 名の計測者 A, B, C, D による CRM スキル計測実験を実施した。

#### 3.2 計測者

実験に参加した CRM スキル計測者 4 名は、いずれも経験豊富な（総飛行時間平均：9125 時間、機長時間平均：3025 時間）機長であり、計測者 A, B, C は所属する航空会社の CRM 部署に従事した経験を有していた。計測者 B は LOFT 教官経験があり、計測者 D は査察操縦士経験があった。また、計測者 A はシナリオ 3 に使用した機種 3 の操縦経験を有していた。計測者には、事前に 2.3~2.5 に述べたような計測のためのガイドラインを示した。

#### 3.3 模擬 LOFT シナリオ

実験には、2.7 に示すシナリオを使用した。

#### 3.4 実験手順

- (1) 3 つのシナリオのビデオを計測者が観察し、提案の計測シートに記入する。シートは、2.4 で述べたように今回提案した Threat ごとに評点するシートと、飛行フェーズごとに評点するシートを用意した。ビデオ観察中において計測者には CRM スキル観測シート<sup>2)</sup>・<sup>3)</sup>（付録 3）に適宜メモを取るよう依頼した。
- (2) (1)終了後、評点の集計を行い 3 つのシナリオの中で比較的ばらつきの少ないシナリオを選択し、評点が分かれた項目について計測者同士の議論を実施した。この議論は、前述したように計測者の間で「True Score<sup>5)</sup>」を導出するためのトライアルである。
- (3) 計測者同士の議論後、ばらつきの大きかったシナリオのビデオを計測者に再

度観測させ、その後 CRM スキル計測シートに記入させ、「True Score<sup>5)</sup>」を導出する議論の有効性を検証した。

### 4. 結果と分析

#### 4.1 全体傾向

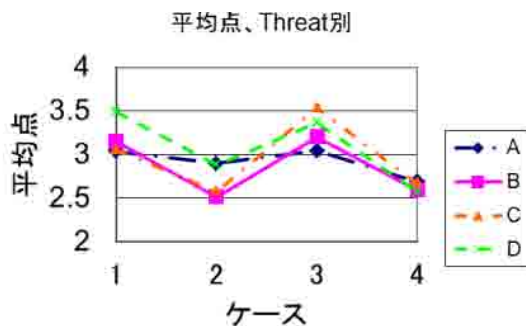
4 名の計測者の評点付けの傾向をみるために、ケース別（ケース 1：シナリオ 1、ケース 2：シナリオ 2、ケース 3：シナリオ 3、ケース 4：シナリオ 2 の 2 回目）に集計した平均点と標準偏差を計算した。平均点のプロットを図 4-1 に、標準偏差を図 4-2 に示す。平均点、標準偏差の算出方法は以下である。なお、評点は順序尺度であるので、等間隔性が保障されておらず、本来であれば平均値や標準偏差を求めることが出来ない。しかし、仮に等間隔と想定すると平均あるいはばらつきに相当する値が得られるので、目安の数値として利用するために、平均値や標準偏差を計算した。

##### (1) 平均点算出方法

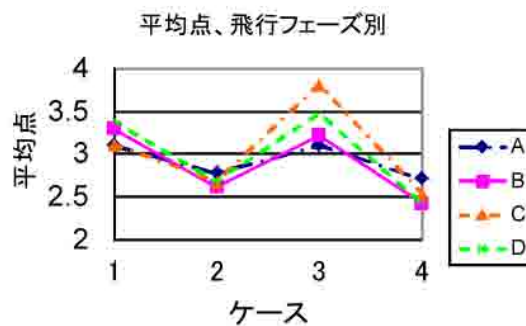
- ①ケース 1 において、計測者 A の評点結果を全て合算。
- ②①の数値、総計測項目数で除算。
- ③残りのケース 2~4 についても同様に①②を繰り返す。
- ④計測者 B~D についても同様に①~③を繰り返す。

##### (2) 標準偏差の算出方法

- ①ケース 1 において、計測者 A の評点結果を全て合算。
- ②先に求めた平均値を使用し、ケース 1 における標準偏差を算出。
- ③残りのケース 2~4 についても同様に、①②を繰り返す。
- ④計測者 B~D についても同様に①~③を繰り返す。

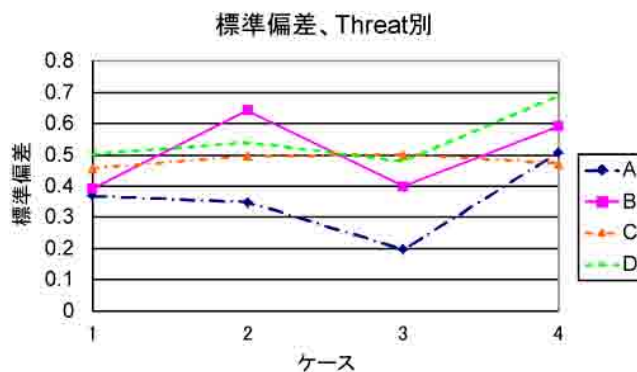


a. Threat 別シート

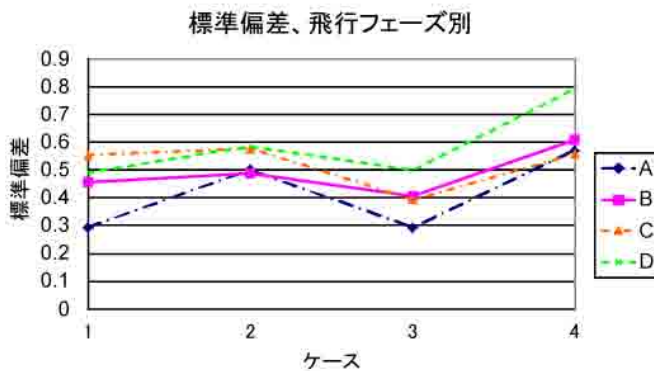


b. 飛行フェーズ別シート

図 4-1 ケース別にみた計測者の平均点推移



a. Threat 別シート



b. 飛行フェーズ別シート

図 4-2 ケース別にみた計測者の標準偏差推移

表 4-1 各計測者で有意な差があったケースの組

計測者	Threat 別		飛行フェーズ別	
	有意なケースの組		有意なケースの組	
A	ケース1	ケース4	ケース1	ケース2
	ケース2	ケース3		ケース4
	ケース3	ケース4	ケース2	ケース3
			ケース3	ケース4
B	ケース1	ケース2	ケース1	ケース2
		ケース4		ケース4
	ケース2	ケース3	ケース2	ケース3
C	ケース1	ケース2	ケース1	ケース2
		ケース4		ケース3
	ケース2	ケース3		ケース4
	ケース3	ケース4	ケース2	ケース3
D	ケース1	ケース2	ケース1	ケース2
		ケース4		ケース4
	ケース2	ケース3	ケース2	ケース3
	ケース3	ケース4	ケース3	ケース4

図 4-1 は計測シート（付録 1, 2 参照）にある「状況の把握・認識の共有」から「意見の相違の解決」までの全計測項目の評点の平均点を示している。ここで A～D は計測者を、横軸はケースを表しており、左から順にケース 1：シナリオ 1、ケース 2：シナリオ 2、ケース 3：シナリオ 3、ケース 4：シナリオ 2 の 2 回目である。すなわち、1 回目の計測実験（3.4(1)の段階）でシナリオ 2 が計測者の間でばらつきがみられたため、3.4(3)の段階でシナリオ 2 の評点を再度行ったことを示している。ここでの平均点とは、2 種類の計測シートごとに集計したものである。

図 4-1 に示された折れ線の変化率より、4 名の計測者の評点付けの傾向が比較的一致していることが分かる。すなわち 4 名ともケース 2 とケース 4 では厳しい評点を、ケース 1 とケース 3 では高い評点付けを行ったと考えられる。さらに、ケース 2 とケース 4 は同じシナリオであるが、ケース 2 と比較してケース 4 の平均点が低くなっている場合が多かった。

4 名の計測者の評点付けの傾向が真に一致しているかを調べるために、一元配置

の ANOVA を用いてケース間の比較を行った。有意な差が存在する場合には多重比較を行った。結果の詳細を付録 4 の 1.1, 1.2 に示す。この結果から、各計測者で平均値の有意な差があったケースを表 4-1 に示す。表 4-1 から、概ね各計測者ともケース 1 に比べて、ケース 2、ケース 4、ケース 2 に比べてケース 3 の評点に有意な差があり、4 名の計測者の評点付けの傾向が一致しているのが分かる。また、各計測者ともケース 2 とケース 4 には有意な差はなかった。

平均点と同様、図 4-2 に示した各ケース毎の標準偏差の折れ線の変化率を見ると、4 名の計測者の評点結果のばらつき傾向が比較的一致している。すなわち計測者 4 名とも、ケース 2 とケース 4 では評点にばらつきが多いことが分かる。さらにケース 2 とケース 4 を比較してみると、同じシナリオでありながら、ケース 4 のばらつきが大きくなっていることが分かる(飛行フェーズ別にみたときの計測者 C、Threat 別にみたときの計測者 B 及び計測者 C を除く)。

## 4.2 提案の計測シートの有効性

図 4-3 に、ケース毎の計測者間の標準偏差平均値（下記，(1)，(2)に導出方法を記載）を示す。図 4-3 より、どのケースにおいても飛行フェーズ別の計測シートと比較して Threat 別の計測シートの計測者間における標準偏差平均値が小さくなっていることが分かる。

そこで、Threat 別と飛行フェーズ別の計測シートにおいて、有意な差があるかを確認するため、標準偏差平均値の独立サンプルの t 検定を実施した。結果の詳細は、付録 4 の 2 に示す。付録の表 2-1~2-4 における等分散性を仮定しない有意確率の結果から、ケース 1 ( $P=0.037<0.05$ )、ケース 3 ( $P=0.000$  以下 $<0.05$ ) には Threat 別と飛行フェーズ別の計測シートで有意な差が認められたが、ケース 2 ( $P=0.100>0.05$ )、ケース 4 ( $P=0.119>0.05$ ) には有意な差は認められなかった。特に 4.1 に述べた評点の平均値が低いケース 2、ケース 4 においては、計測シートの種別で差はなかった。

- ケース別にみた標準偏差の平均値算出方法

### A. Threat 別

- ①ケース 1 について、計測項目の標準偏差  $SD_1 \sim SD_n$  を導出（計測項目  $14 \times$  Threat の数）。
- ② $SD_1 \sim SD_n$  の平均値  $SD_{11}$  を導出。
- ③ケース 2 からケース 4 についても、同様に標準偏差の平均値  $SD_{22}$ 、 $SD_{33}$ 、 $SD_{44}$  を導出する。

### B. 飛行フェーズ別

- ①ケース 1 について、計測項目の標準偏差  $SD_1 \sim SD_n$  を導出（計測項目  $14 \times$  飛行フェーズ（Pre departure/Taxi Out, Take off, Cruise, Descend / Landing）の数）。
- ② $SD_1 \sim SD_n$  の平均値  $SD_{11}$  を導出。
- ③ケース 2 からケース 4 についても、同様に標準偏差の平均値  $SD_{22}$ 、 $SD_{33}$ 、 $SD_{44}$  を導出する。

今回の 2 種類の計測シートについて、計測者のコメントから、主な長所と短所は以下の通りであった。

#### (1) 飛行フェーズ別

- ・ひとつの飛行フェーズで複数回スキルが観察される場合もあり、複数の評点結果が相殺される危険性がある。

- ・飛行フェーズを通して全体評価が可能。

#### (2) Threat 別

- ・詳細に評点が可能。

- ・Threat による後々まで及ぶ影響を評価できない。

Threat は存在しなくても、CRM スキルが観察された／観察されない場合、Threat に関係付けることができず、シートへの記入が難しい。

## 4.3 計測者同士の議論の有効性

前述したように、ケース 4 を行う前に、ケース 1 及びケース 3 の評点結果等を議題に計測者同士で「True Score<sup>5)</sup>」導出のための議論を行わせた。その議論が終了した後、ケース 4 を実施した。ケース 4 はケース 2 と同じ LOFT に対するクルーの CRM スキルが評点対象であるためケース 2 とケース 4 の結果の違いは、計測者同士の議論と、計測者にとって十分精通したシナリオであることに起因する。

ケース 2 と 4 における標準偏差の平均値を図 4-4 に示す。この図から、Threat 別、飛行フェーズ別のシートとも、ケース 4 の方が、ばらつきが大きくなっていることが分かる。

そこで、ケース 2 とケース 4 において、有意な差があるかを確認するため、標準偏差平均値の t 検定を実施した。ケース 2 とケース 4 は、同じシナリオであり、計測者同士の議論前と議論後を比較しているので、各観測値が両方の母集団平均に対応していると考えられる。そこで、対応のある t 検定を用いて有意差検定を行った。結果の詳細は、付録 4 の 3 に示す。この結果から、Threat 別 ( $P=0.399>0.05$ )、飛行フェーズ別 ( $P=0.382>0.05$ ) とともに、ケース 2 とケース 4 で有意な差は認められなかった。この結果は各計測項目に対して、標準偏差を導出し、さらにその標準偏差の平均値を出しているため、この平均値の比較だけでは、ケース 2 とケース 4 の具体的な差を明らかにできない可能性がある。

そこで、どのような変化が生じていたかを明らかにするために評点の割合の変



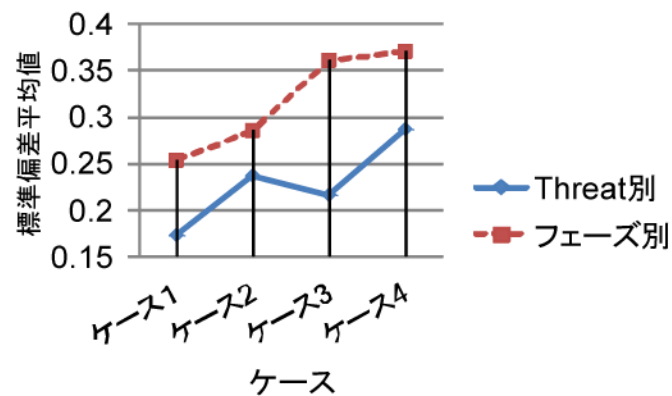


図 4-3 ケース別にみた標準偏差の平均値

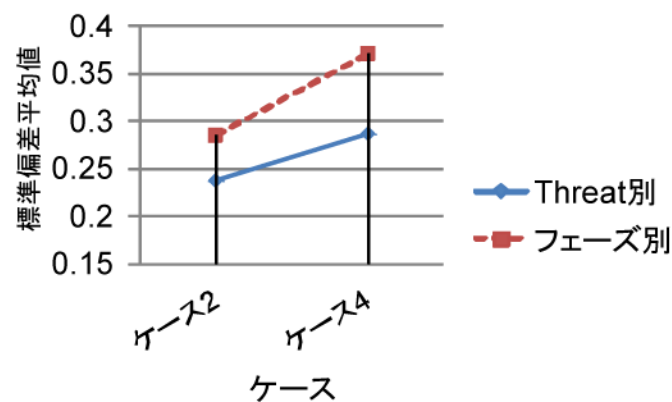


図 4-4 ケース 2 と 4 の標準偏差平均値

化を図 4-5 に、さらに細かい数値を表 4-2 にまとめた。

図 4-5 はケース 2 とケース 4 を比較して、「1 点」、「2 点」、「3 点」、「4 点」の割合、そしてなにも評点しなかった空欄の割合が何%存在したかを表している。これらの図から、ケース 2 の方はどちらのシートとも、空欄の割合以外は 3 点の占める割合がほとんどであり、平均に偏った評点をしていたことが分かる。ケース 4 の方は、3 点の割合が減って、2 点、1 点の割合が増加して、めりはりのある評点となっていることが分かる。

表 4-2 から読み取れることをまとめる。  
飛行フェーズ別計測シートの結果から読み取ることができる事柄としては、  
・計測者全員の平均点がケース 2 に比べてケース 4 で低下した。ただし、表 4-1

に示したように、ケース 2 とケース 4 の平均点には有意な差はない。

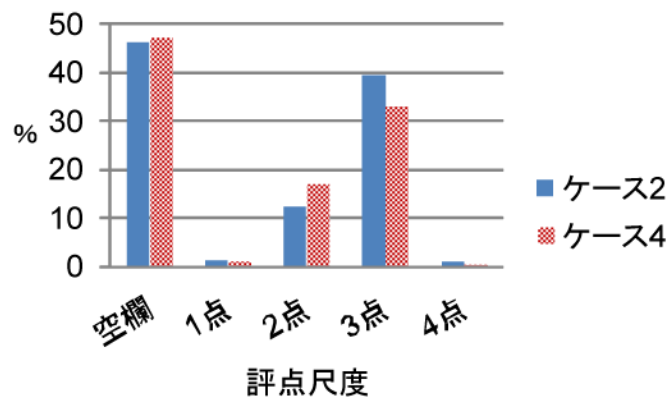
- ・4 名のうち 3 名の標準偏差が増大。
- ・4 名のうち 3 名の「1 点」の数が増加。
- ・計測者全員の「3 点」の数が増加。
- ・計測者全員において「空欄数」が増加。

一方、Threat 別計測シートからは

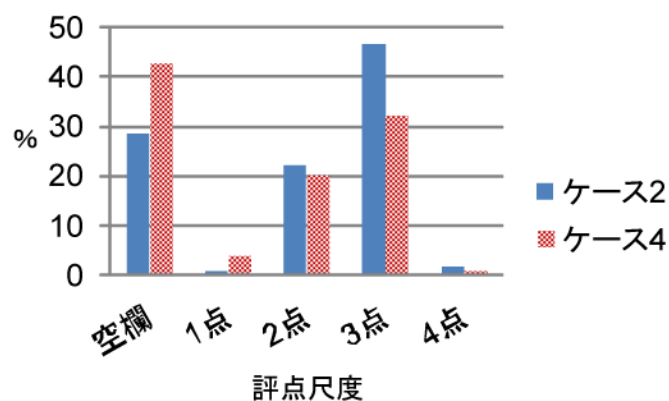
- ・4 名のうち 2 名の平均点がケース 2 に比べてケース 4 で低下した。ただし、表 4-1 に示したように、ケース 2 とケース 4 の平均点には有意な差はない。
- ・4 名のうち 2 名の標準偏差が増大。
- ・4 名のうち 3 名の「2 点」の数が増加等が挙げられる。

どちらの計測シートの結果においてもケース 4 において「3 点」の数が増加し、「1 点」「2 点」の数が増加している。





a. Threat 別



b. 飛行フェーズ別

図 4-5 計測者の評点割合の平均値

表 4-2 ケース 2 とケース 4 における各評点の変化

a. 飛行フェーズ別

		A	B	C	D
平均点	2	2.77	2.61	2.67	2.67
	4	2.71	2.42	2.53	2.42
標準偏差	2	0.505	0.495	0.586	0.595
	4	0.579	0.620	0.567	0.807
1点の数	2	2	0	0	0
	4	2	1	1	5
2点の数	2	8	15	14	13
	4	6	17	13	9
3点の数	2	43	23	20	18
	4	26	12	18	16
4点の数	2	0	0	2	2
	4	0	1	0	1
空欄数	2	3	18	20	23
	4	22	25	24	25

表 4-2 ケース 2 とケース 4 における各評点の変化

		b. Threat 別			
		A	B	C	D
平均点	2	2.90	2.47	2.53	2.88
	4	2.69	2.59	2.67	2.57
標準偏差	2	0.369	0.667	0.546	0.516
	4	0.512	0.599	0.476	0.695
1 点の数	2	1	3	0	0
	4	1	1	0	3
2 点の数	2	5	12	18	7
	4	13	14	16	15
3 点の数	2	63	22	24	23
	4	34	21	32	24
4 点の数	2	0	0	0	3
	4	0	1	0	2
空欄数	2	15	47	42	51
	4	36	47	36	40

以下に、計測者同士の議論の主な結果をまとめる。

- ・ 評点が修正された。例えば、1 点を付けた計測者が 2 点に修正し、4 点をつけた計測者が 2 点に修正した。
- ・ 最初の評点では、全体的印象で評点する傾向だったが、議論後の評価ではより客観的に評点出来たと言う計測者が存在した。すなわち計測者は、第一印象が以後の評点に影響するといった偏った評点をしがちなので練習が必要である。
- ・ 最初の評点では、可否の判定の観点で評点したが、議論後、訓練の観点（再訓練が必要か否か）で評点するようになった計測者が存在した。
- ・ 計測者同士の議論において、1 点が 2 点、4 点が 2 点のように収束していく傾向はあるものの 2 点と 3 点の間で「True Score<sup>5)</sup>」を出すのは難しい（どちらかの点になる）。
- ・ 計測者となるには、当該機種の乗員資格は必ずしも必須ではないが、少なくとも当該機種に関する SOP の知識は必須である。
- ・ 評点のばらつきが改善されたとしても、各計測者によって着目する CRM スキルが異なることがある。そのキャリブ

レーションが課題である。そもそもキャリブレーションは可能なのかという疑問もある。

- ・ 同一のシナリオを 2 回観察した後の評点においても、必ずしも計測者同士の標準化がなされたとは限らないことに注意する。行動を注意深く観測できるようになった、あるいは評点に慣れたために、1 回目と 2 回目の評点が異なることがある。

#### 4.4 計測者の評点傾向一致の度合い

各ケースにおいて計測者の評点傾向一致の度合いをみるために、ケース別に、計測者の計測項目評点の平均値及び、標準偏差を計算した。その集計結果を表 4-3 に示す。

##### (1) 平均点の算出方法

- ① ケース 1 において、計測者 A～D 全員の評点結果をそのまま合算。
- ② ①の数値を、総計測項目数で除算。
- ③ 残りのケース 2～4 についても同様に、①②を繰り返す。

##### (2) 標準偏差の算出方法

- ① ケース 1 において、計測者 A～D 全員の評点結果をそのまま合算。

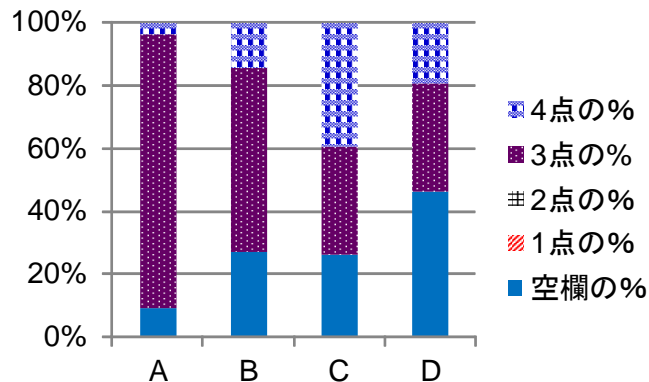
- ②先に求めた平均値を使用し、ケース 1 における全評点の標準偏差を算出。  
 ③残りのケース 2～4 についても同様に、  
 ①②を繰り返す。

表 4-3 から、Threat 別、飛行フェーズ別とも、平均点はケース 3 が最も高く、逆にケース 4 が最も低い結果であった。標準偏差はケース 4 が最も高く、ケース 1 が最も低い結果であった。

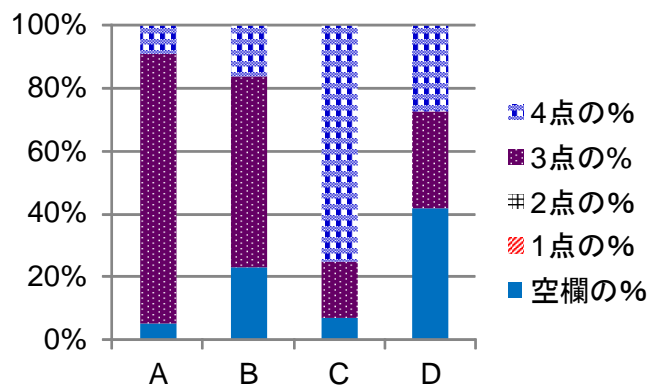
ケース毎に有意差があるかどうかを調べるため、一元配置の ANOVA を実施した。有意な差が存在する場合には多重比較を行った。結果の詳細を付録 4 の 4.1, 4.2 に示す。この結果から、Threat 別、飛行フェーズシートとも有意な差が認められた。両シートとも、ケース 2 とケース 4 の組以外は、有意な差が認められた。

表 4-3 ケース別にみた平均点・標準偏差

ケース	シート種類	平均点	標準偏差
1	飛行フェーズ別	3.17	0.467
	Threat 別	3.12	0.432
2	飛行フェーズ別	2.69	0.540
	Threat 別	2.71	0.544
3	飛行フェーズ別	3.39	0.490
	Threat 別	3.26	0.441
4	飛行フェーズ別	2.52	0.652
	Threat 別	2.63	0.570

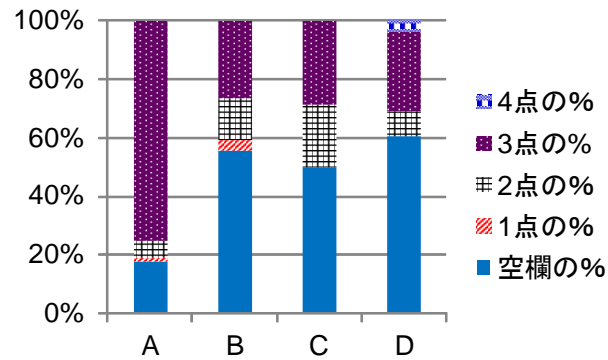


a. Threat 別

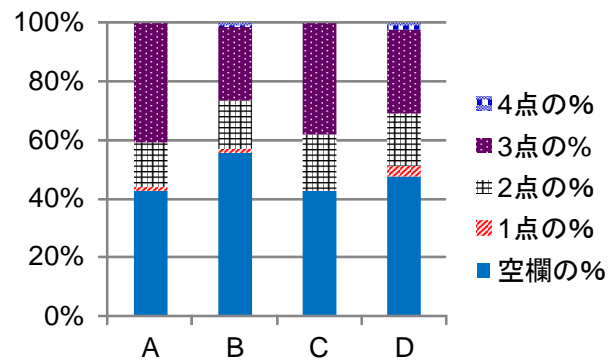


b. 飛行フェーズ別

図 4-6 CRM スキル計測シート集計\_ケース 3

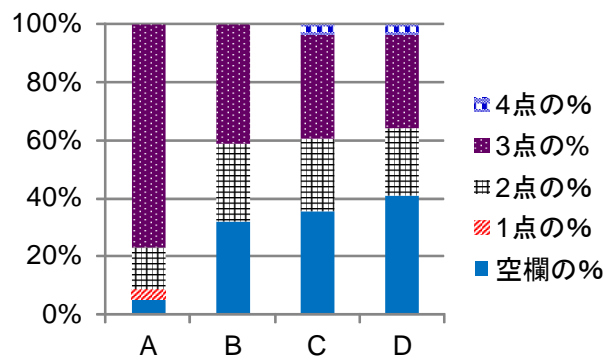


a. ケース 2

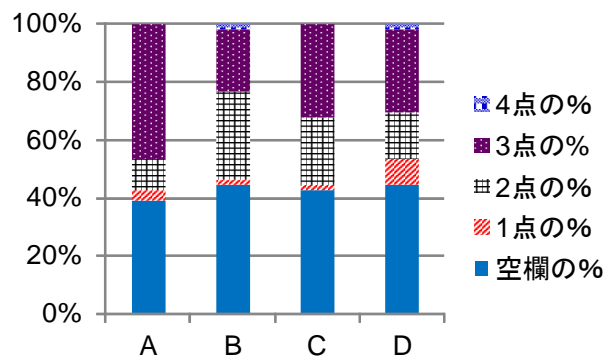


b. ケース 4

図 4-7 CRM スキル計測シート集計\_Threat 別



a. ケース 2



b. ケース 4

図 4-8 CRM スキル計測シート\_飛行フェーズ別

表 4-4 計測者別にみた平均点及び標準偏差

計測者	シート種類	平均点	標準偏差
A	飛行フェーズ別	2.94	0.452
	Threat 別	2.94	0.378
B	飛行フェーズ別	2.90	0.614
	Threat 別	2.93	0.580
C	飛行フェーズ別	3.12	0.735
	Threat 別	2.99	0.628
D	飛行フェーズ別	2.98	0.759
	Threat 別	3.03	0.666

ここで、平均点が最も高かったケース 3 では 4 名の計測者の標準偏差がケース 1 の次に小さい値になっている。つまり、ケース 3 を評点付けしたとき、全ての計測者が比較的高い評点を付けたと推察される。このことは、図 4-6 に示すケース 3 の CRM スキル計測シートの集計結果からも明らかである。図 4-6 から、どの計測者においても「2 点」，「1 点」は存在せず、ほとんど「3 点」，「4 点」をつけていたのが分かる。

一方、標準偏差をみると、最もこれが大きかったケース 4 では、平均点が 2.52 と小さい値になっている。つまり、計測者の評点付けは厳しいものであり、評点付けの傾向も、計測項目間あるいは計測者間で差があったことがうかがえ、これは図 4-7、図 4-8 に示す結果からも明らかである。

これらの事から、計測者間での評点の平均値が低い場合は、ばらつきが大きい傾向であることが分かる。

#### 4.5 各計測者の評点傾向

各計測者の評点傾向（例：高い点を付けやすい、平均値に近い点数を付けがち）を見るために、計測者別平均点及び、標準偏差を計算した。計測者別にみた平均点及び標準偏差の平均値を表 4-4 に、各計測者のケースごとの評点の割合を図 4-9 に示す。

##### (1) 平均点の算出方法

①計測者 A において、ケース 1～4 の評点

結果をそのまま合算。

②①の数値を、総計測項目数で除算。

③計測者 B～D についても同様に、①②を繰り返す。

##### (2) 標準偏差の算出方法

①計測者 A において、ケース 1～4 の評点結果をそのまま合算。

②先に求めた平均値を使用し、計測者 A における全評点の標準偏差を算出。

③計測者 B～D についても同様に、①②を繰り返す。

表 4-4 から平均点は計測者 C(飛行フェーズ別)の結果が最も高く、計測者 B(飛行フェーズ別)が最も低かった。

計測者毎に有意差があるかどうかを調べるため、一元配置の ANOVA を実施した。有意な差が存在する場合には多重比較を行った。結果の詳細を付録 4 の 5.1, 5.2 に示す。この結果から、Threat 別シートでは有意な差は認められなかった。飛行フェーズ別シートでは、計測者 C は、他計測者 A ( $P = 0.034 < 0.05$ )、B ( $P = 0.028 < 0.05$ )、と有意な差があったが、計測者 A, B 間 ( $P = 0.944 > 0.05$ )、計測者 B, D 間 ( $P = 0.834 > 0.05$ ) には有意な差は存在しなかった。

標準偏差は、計測者 D(飛行フェーズ別)の結果が最も高く、計測者 A(Threat 別)が最も低かった。

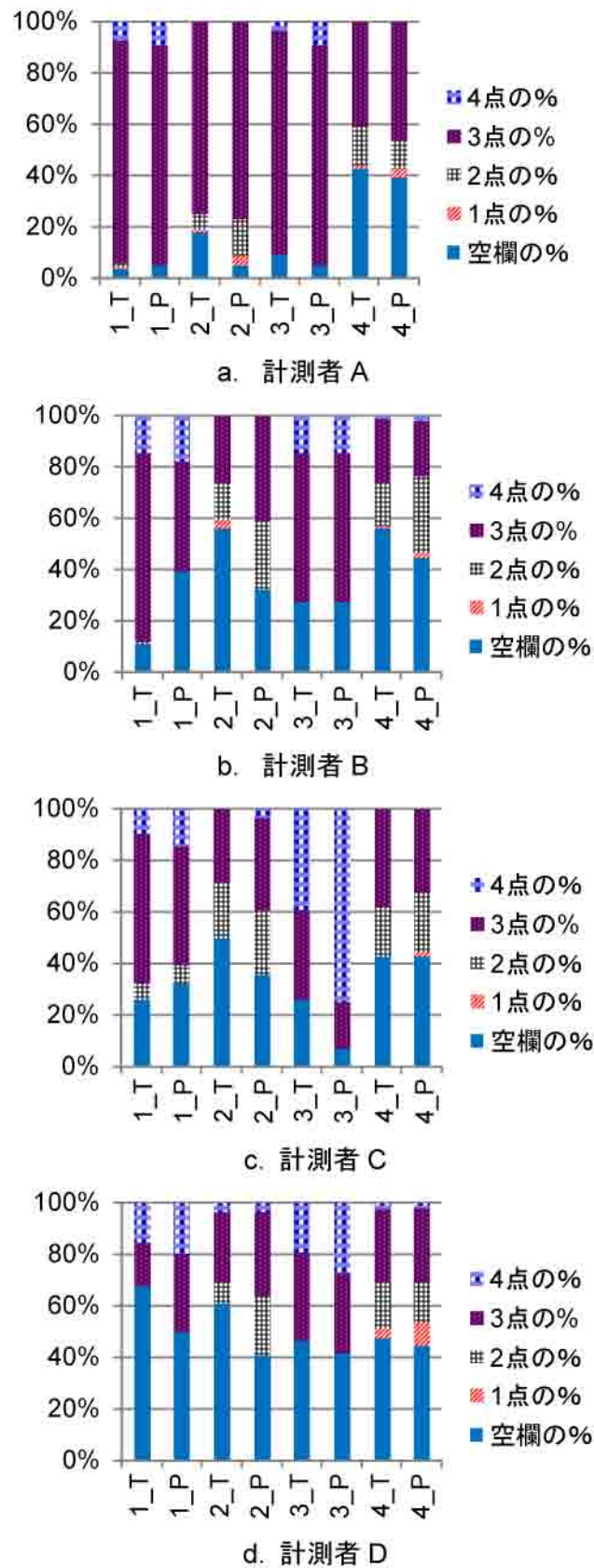


図 4-9 CRM スキル計測シート集計\_計測者別

注: 1\_T~4\_T : ケース1~ケース4のThreatごと測定

1\_P~4\_P : ケース1~ケース4のPhaseごと測定

計測者 C は、平均点が高いことから一見、どのシナリオにおいても甘い評点付けを行ったかのように見えるが、標準偏差が大きいことから、評点にばらつきがあり、高い評点と低い評点とが混在していることが推察され、これは図 4-9-c から明らかである。

計測者 A の標準偏差が小さいのは、図 4-9-a から、他計測者に比べて 3 点の割合が大きいからであり、他計測者に比べて、平均点を評点しやすい傾向であることが推察される。しかし、計測者 A のケース 2 とケース 4 を比べると、同じシナリオでありながら、ケース 4 において 3 点の割合が減っていた。

#### 4.6 計測者間のばらつきが大きい CRM スキル計測項目

4 名の計測者の標準偏差が大きい CRM スキル計測項目の上位 10 項目を表 4-5 に示す。表 4-5-a, b はそれぞれ、飛行フェーズ別計測シート、Threat 別計測シートの結果である。右端の欄の数字が、ケース 1~4 における各 CRM スキル計測項目の標準偏差の平均値である。例えば、計測項目のひとつである「状況認識マネジメント」の「状況の把握・認識の共有」は、飛行フェーズごと、もしくは Threat ごとに評点欄があるので、同じ計測項目の標準偏差を全て合計した。さらに、それをシナリオ 4 本分に合計し、計測者 4 名分を合計し、平均値を算出した。

##### (1) 計測項目ごとの標準偏差算出方法(飛行フェーズ別シートの場合)

①ケース 1 における、1 つの飛行フェーズの 1 つの計測項目において、計測者 4 名の評点の標準偏差 SD\_1 を算出。

例：「Pre departure / Taxi-out」/計測項目「状況の把握・認識の共有」において、計測者 4 名の評点の標準偏差を算出。

②残りのケース 2~4 においても同様に①を実施。ここで、全てのケース 1 から 4 において、1 つの飛行フェーズの 1 つの計測項目の標準偏差 (SD\_1, SD\_2, SD\_3, SD\_4) が算出されたことになる。

例：ケース 1~4 それぞれにおいて、「Pre departure / Taxi-out」/計測項目「状況の把握・認識の共有」の計測者 4 名の評点の標準偏差を算出

③②にて求めた標準偏差の合計値 (SD\_1 + SD\_2 + SD\_3 + SD\_4) から平均値 (Average\_1) を算出。

ここで、1 つの飛行フェーズにおける 1 つの計測項目の標準偏差の平均値が算出されたことになる。

④他の計測項目においても同様に①~③を実施。

ここで、1 つの飛行フェーズにおける全ての計測指標の標準偏差の平均値 (Average\_1~Average\_n: n=12, 12 は計測項目の数) が算出されたことになる。

⑤他の飛行フェーズ「Takeoff/Climb」, 「Cruise」, 「Descend/Approach/Land」においても、①~④を実施。

これで、全ての飛行フェーズにおける全ての計測指標の標準偏差の平均値が算出されたことになる。

導出された平均値の個数 = 飛行フェーズの数 × (Average\_1~Average\_n: n=12)。

##### (2) 計測項目ごとの標準偏差算出方法(Threat 別シートの場合)

①ケース 1 における、1 つの Threat の 1 つの計測項目において、計測者 4 名の評点の標準偏差 SD\_1 を算出。

例：Threat「乗客にけが」/計測指標「状況の把握・認識の共有」において、計測者 4 名の評点の標準偏差を算出。

②残りのケース 2~4 においても同様に①を実施。

ここで、全てのケース 1~4 において、1 つの Threat の一つの計測項目の標準偏差 (SD\_1, SD\_2, SD\_3, SD\_4) が算出されたことになる。

③②にて求めた標準偏差の合計値 (SD\_1 + SD\_2 + SD\_3 + SD\_4) から平均値 (Average\_1) を算出。

ここで、1 つの Threat における 1 つの計測項目の標準偏差の平均値が算出されたことになる。



表 4-5 標準偏差の大きかった CRM スキル計測項目

## a. 飛行フェーズ別

DAL	状況認識マネージメント	状況の把握・認識の共有	0.584
DAL	チーム形成・維持	チーム活動に適した雰囲気・環境作り	0.579
DAL	コミュニケーション	安全への主張	0.565
TOC	状況認識マネージメント	状況の把握・認識の共有	0.538
CRZ	ワークロードマネージメント	タスクの配分	0.521
CRZ	意思決定	解決策の選択	0.493
TOC	チーム形成・維持	業務の主体的遂行	0.486
DAL	状況認識マネージメント	警戒・予測	0.476
CRZ	状況認識マネージメント	状況の把握・認識の共有	0.459
DAL	コミュニケーション	ブリーフィング	0.454
TOC	状況認識マネージメント	警戒・予測	0.452

## b. Threat 別

TOC	チーム形成・維持	業務の主体的遂行	0.457
TOC	コミュニケーション	ブリーフィング	0.456
TOC	状況認識マネージメント	問題点の分析	0.419
TOC	状況認識マネージメント	警戒・予測	0.391
TOC	ワークロードマネージメント	タスクの配分	0.390
TOC	チーム形成・維持	チーム活動に適した雰囲気・環境作り	0.385
CRZ	コミュニケーション	情報の伝達と確認	0.357
TOC	コミュニケーション	情報の伝達と確認	0.352
DAL	状況認識マネージメント	問題点の分析	0.317
DAL	チーム形成・維持	業務の主体的遂行	0.317
DAL	ワークロードマネージメント	プランニング・優先順位付け	0.313

注 TOC: Take off / Climb, CRZ: Cruise, DAL: Descend / Approach / Land

④他の計測項目においても同様に①~③を実施。

ここで、1 つの Threat における全ての計測項目の標準偏差の平均値 (Average<sub>1</sub>~Average<sub>n</sub>: n=12, 12 は計測項目の数) が算出されたことになる。

⑤他の Threat においても、①~④を実施。

これで、全ての Threat における全ての計測項目の標準偏差の平均値が算出されたことになる。

導出された平均値の個数 = Threat 数 × (Average<sub>1</sub>~Average<sub>n</sub>: n=12) 。

表 4-5-a より、飛行フェーズ別で評点すると

- ・飛行フェーズの「Descend / Approach / Land」(上位 3 位を占め、抽出された数も多い)

- ・CRM スキルの「状況認識マネージメント」(上位 1 位を占め、抽出された数も

多い)

において、標準偏差が大きい、つまり評点にばらつきが多いことが分かる。

また、表 4-5-b より Threat 別で評点すると・飛行フェーズの「Takeoff/Climb」(ほとんどが、この飛行フェーズ)において、標準偏差が大きい、つまり評点にばらつきが多いことが分かる。Threat ごとでは、最もばらついているのは「チーム形成・維持」であるが、抽出された数が比較的多いのは「状況認識マネージメント」と「コミュニケーション」であった。

表 4-5-a, b を比較すると、飛行フェーズ別の標準偏差値が Threat 別のそれに比べて、若干大きい傾向にあった。

飛行フェーズ別でみると「Descend / Approach / Land」, Threat 別でみると「Take Off / Climb」でばらつきが大きくなっているのは興味深い結果であるが、そ



表 4-6 空欄が目立った CRM スキル計測項目(飛行フェーズ別)

3.25	Predeparture/Taxiout	コミュニケーション	安全への主張
3.25	Predeparture/Taxiout	チーム形成・維持	意見の相違の解決
3.25	Takeoff/Climb	チーム形成・維持	意見の相違の解決
2.75	Predeparture/Taxiout	意思決定	決定・行動のレビュー
2.75	Takeoff/Climb	コミュニケーション	安全への主張
2.50 2.50 2.50 2.50 2.50	Predeparture/Taxiout	意思決定	解決策の選択
	Predeparture/Taxiout	ワークロードマネジメント	タスクの配分
	Takeoff/Climb	意思決定	決定・行動のレビュー
	Cruise	コミュニケーション	安全への主張
	Cruise	チーム形成・維持	意見の相違の解決
	Descend/Approach/Land	チーム形成・維持	意見の相違の解決

表 4-7 空欄が目立った CRM スキル計測項目(Threat 別)

3.27	Cruise	チーム形成・維持	意見の相違の解決
3.00 3.00 3.00 3.00	Predeparture/Taxiout	状況認識マネジメント	問題点の分析
	Predeparture/Taxiout	コミュニケーション	ブリーフィング
	Predeparture/Taxiout	コミュニケーション	安全への主張
	Predeparture/Taxiout	チーム形成・維持	チーム活動に適した雰囲気・環境作り
2.73	Cruise	コミュニケーション	安全への主張
2.67	Takeoff/Climb	コミュニケーション	安全への主張
2.55	Cruise	意思決定	決定・行動のレビュー
2.36	Cruise	チーム形成・維持	チーム活動に適した雰囲気・環境作り
2.33	Takeoff/Climb	チーム形成・維持	意見の相違の解決
2.33	Descend/Approach/Land	コミュニケーション	安全への主張

の理由を示す解析結果や計測者のコメントは抽出されなかった。

#### 4.7 空欄が目立つ CRM スキル計測項目

空欄が多く見られた CRM スキル計測項目の上位 10 項目を表 4-6 および表 4-7 に示す。左端の欄の数字が、実験ケース 1～4 における空欄数の平均値である。例えば、「コミュニケーション」の「安全への主張」では、平均値が 3.25 であることから、4 回の実験ケースにおいて  $3.25 \times 4 = 13$  箇所が空欄であったことを意味する。

ここで、評価者 4 名が全員揃って空欄としていた箇所は「観測できなかった」という結果で 4 名が同意していたものとみなし、対象からは外している。

表より、空欄は

- ・飛行フェーズ別シート of 「Predeparture/Taxiout」
  - ・CRM スキルの「コミュニケーション」
  - 「チーム形成・維持」
  - ・CRM スキル計測項目の「意見の相違の解決」「安全への主張」
- において多く発生する傾向があることが分かる。

表 4-8 実験ケース 3 における操縦経験有無による影響

		A	B	C	D
平均点	飛行フェーズ別	3.09	3.21	3.81	3.47
	Threat 別	3.04	3.20	3.53	3.36
標準偏差	飛行フェーズ別	0.295	0.412	0.398	0.507
	Threat 別	0.196	0.401	0.503	0.484

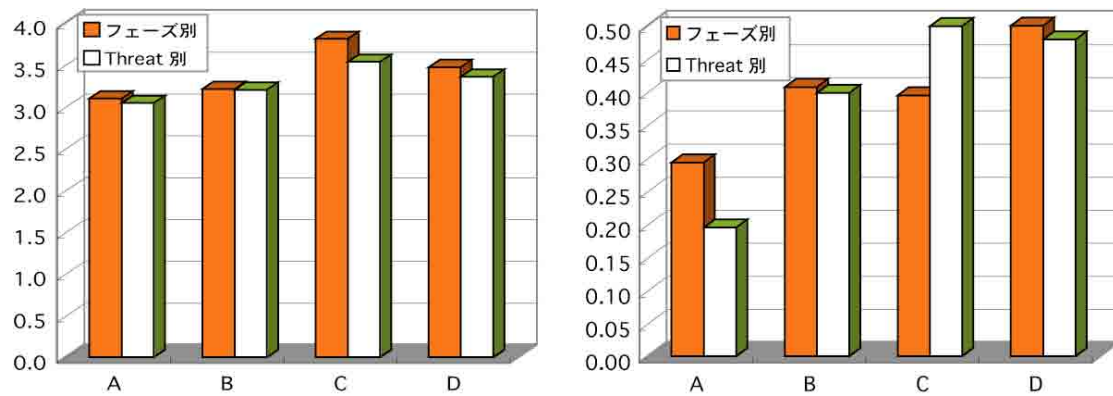


図 4-10 実験ケース 3 における各評価者の平均値(左)と標準偏差(右)

表 4-9 実験ケース 2 とケース 4 における各評点の変化(Threat 別)

ケース	Threat 数	飛行時間 (分)	1 分当たりの Threat 数	4 名の計測者平均値		標準偏差
1	6	90	0.067	飛行フェーズ別	3.17	0.467
				Threat 別	3.12	0.432
2	6	100	0.060	飛行フェーズ別	2.69	0.540
				Threat 別	2.71	0.544
3	6	80	0.075	飛行フェーズ別	3.39	0.490
				Threat 別	3.26	0.441
4	6	100	0.060	飛行フェーズ別	2.52	0.652
				Threat 別	2.63	0.570

表 4-10 Threat の種類と標準偏差

Threat 種類 ケース	Aircraft	Departure / Arrival	Cabin	Operational	ATC	4 名の標準偏差 平均値		標準偏差
1	3	2	1	0	0	飛行フェーズ別	3.17	0.467
						Threat 別	3.12	0.432
2	3	1	1	1	0	飛行フェーズ別	2.69	0.540
						Threat 別	2.71	0.544
3	1	2	1	1	1	飛行フェーズ別	3.39	0.490
						Threat 別	3.26	0.441
4	3	1	1	1	0	飛行フェーズ別	2.52	0.652
						Threat 別	2.63	0.570

#### 4.8 当該機種の操縦経験有無による相違

今回の CRM スキル計測実験では、計測者 A のみ、実験ケース 3 である機種 3 の経験を有していた。そこで、実験ケース 3 において、計測者 A と B～D の結果を比較した。

表 4-8 および図 4-10 より、操縦経験を有する計測者 A は、平均点および標準偏差において、他の 3 名よりも評点が厳しく、かつ点数のばらつきが少ないことが分かる。特に、標準偏差に関しては差が顕著である。計測者 A が他計測者と有意差があるかどうかを調べるため、一元配置の ANOVA を実施した。有意な差が存在する場合には多重比較を行った。結果の詳細を付録 4 の 6.1, 6.2 に示す。この結果から、Threat 別シートでは、計測者 A は、他計測者 B ( $P=0.031<0.05$ )、C ( $P=0.000$  以下  $<0.05$ )、D ( $P=0.001<0.05$ ) と有意な差が認められ、飛行フェーズ別シートでは、計測者 A は、計測者 C ( $P=0.000$  以下  $<0.05$ )、計測者 D ( $P=0.002<0.05$ ) と有意な差は認められた。

#### 4.9 LOFT シナリオの有効性

##### (1) Threat の数との関連

シナリオに含まれる Threat の数によって、計測者の評点のばらつき度合いが影響を受けるかどうかを調べるため、表 4-9 に対応を示す。ここでの Threat の数とは、予め組み込まれていた Threat の数であり、計測者によって後から抽出された Threat は含めていない。

表 4-9 に示される平均点、標準偏差は表 4-3 に示した値と同じである。

表 4-9 から、Threat の数が同じであるにも関わらず、平均点やばらつきの大きさに差があることが分かる。4.4 に述べたように、ケース 2 とケース 4 以外は有意な差が存在した。この結果からは、Threat の数が、評点にどのように影響するかを明らかにすることはできなかった。

##### (2) Threat の種類との関連

シナリオに含まれる Threat の種類によって、計測者の評点のばらつき度合いが影響を受けるかどうかを調べるため、表 4-10 に対応を示す。表 4-10 から様々な Threat が満遍なく入っているシナリオ 3

(ケース 3) の 4 名の計測者の平均点が高い事が分かる。ただし、シナリオ 3 のクルーで他シナリオを比較した訳ではないので、様々な Threat が入っていると平均点が高くなると一概にはいえない。

Threat の種類の他に、どのような種類の Threat をどのタイミングで入れるかも重要であると思われる。シナリオ 1, 2 の両シナリオには Aircraft Threat が 3 つ入っているが、シナリオ 2 の場合は 2.7 でも述べたように Aircraft Threat が立て続けに入っており、シナリオ 1 の方は Aircraft Threat の間に他の Threat を挿入している。

Aircraft Threat が立て続けに入っているシナリオ 2 (ケース 2, ケース 4: ケース 2, 4 の平均点に有意差なし) は、平均点が低く、ばらつきも大きかった。

#### (3) シナリオ評価

表 4-11, 各シナリオに対する被計測者及び LOFT 教官の評点とコメントを示す。LOFT 教官は、本 CRM スキル計測シートを用いた CRM スキル計測は実施せず、通常の LOFT のように、シミュレータでのイベントの操作や、ブリーフィング、デブリーフィングを担当した。評点は、LOFT シナリオが、実運航に即していたかの問いに対して、1 点：即していない、2 点：あまり即していない、3 点：即していた、4 点：大変即していたの 4 段階とした。

表 4-11 から、概ね (シナリオ 3 の教官 C 以外)、シナリオは実運航に即していたとの評点だった。この結果からは、実運航に即しているシナリオが計測者の評点にどのように影響するかを明らかにすることはできなかった。

シナリオ 2 の FO B, シナリオ 3 の CAP C のコメントから、テクニカルなイベント (Aircraft Threat) が続くことは、CRM スキル訓練というより、プロセジャ訓練に重きがおかれる可能性があることを示唆していることが分かる。

#### 4.10 チームの総合点

今回は総合点をどのような基準で評点するかの指標を特に設けずに計測者に試験的に評点付けさせたが、図 4-11 から、チームの総合点は計測者間で比較的一致する傾向にあった。チームの総合点については、Threat 別、飛行フェーズ別計測

シートとも同じ点数であった。ケース 3 は計測者全員が 4 点をつけており、ケース 2 と 4 では、計測者 D 以外は 2 点、ケース 1 では計測者 C 以外は 4 点と一致していた。ケース 2 とケース 4 で各計測者において評点が変わることはなかった。

表 4-11 LOFT シナリオが実際の運航に即しているかどうかのアンケート結果

	被験者	評点	コメント
シナリオ 1	CAP A	3	なし
	FO A	3	ハリーアップシンドロームに入りやすい状況の中での判断は勉強になった。
	教官 A	4	なし
シナリオ 2	CAP B	3	なし
	FO B	3	いずれのトラブルも、実運航で発生しうるものと思われましたが、トラブルの数が多く、クルー間でいろいろ相談するというよりは、トラブルに対するプロセジャに終始してしまったように思います。
	教官 B	3	なし
シナリオ 3	CAP C	3	イベントの数と内容が 現実的 であり、受訓者に CRM スキル を発揮させる場面をうまく設定した良いシナリオであった。テクニカルなトラブルが次々に発生するようなシナリオではプロセジャ訓練 になってしまうが、このシナリオは プロセジャ (チェックリスト) が無い故障を 1 つだけ発生させ、クルーに考えさせながら、空港閉鎖やけが人発生によるストレスをかけるなど、演練すべき CRM スキルをベースに イベント を組み込んであり、効果的だった。
	FO C	3	乗務中にストレスを感じるのはトラブル、Threat 等が重なったときであり、実運航でも起こりうるものと思います。
	教官 C	2	空港選択がある程度限られた中でのシナリオはやや無理がある（仕方のないことですが・・・）。クルーの選択肢の巾をなくすことが場合によっては Threat となりうる。

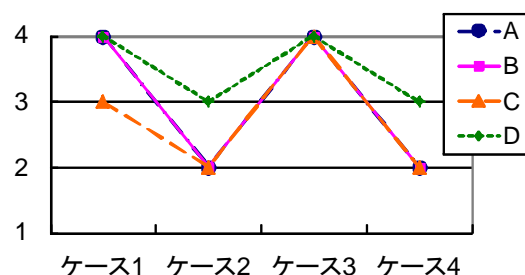


図 4-11 チーム総合点比較

## 5. 考察

CRM スキル計測手法は、「どのような Threat や Error に対して、何のスキルが欠如しているか、それを補う為にどのような訓練が必要か？」を分析出来ることが最終目標である。データ取得のためのツールである「CRM スキル計測シート」、信頼するデータを得るための「計測者の標準化」、CRM スキルを有効に発揮させるための「LOFT シナリオ開発」が技術課題となる。以下に、「CRM スキル計測シート」、「計測者の標準化」、「LOFT シナリオ開発」の観点で考察する。

### 5.1 CRM スキル計測シート

#### (1) Threat 別 vs. 飛行フェーズ別計測シート

4.2 に述べたように、飛行フェーズ別の計測シートと比較して Threat 別の計測シートの標準偏差平均値が小さくなったケース（ケース 1, ケース 3）が存在した。一方で、ケース 2, ケース 4 は、Threat 別計測シートで標準偏差平均値が小さくなったものの、有意な差ではなかった。

また、4.6 の表 4-5-a, b から、標準偏差の大きかった CRM スキル項目についても、飛行フェーズ別の標準偏差値が Threat 別のそれに比べて、若干大きい傾向にあった。これらの結果から、Threat 別に評点付けを行うことで、計測者間のばらつきがある程度は抑えられたと考えられる。さらに、4.5 に述べたように、Threat 別シートでは各計測者間の有意な差は認められなかったことから、Threat 別の計測シートを用いることで、偏った評点を軽減することができ、評点傾向を一致させることができると考えられる。

飛行フェーズ別の標準偏差値が Threat 別に比べて大きい傾向である理由は、計測者のコメントから、飛行フェーズごとに評点すると同じ飛行フェーズで「良い行動」、「悪い行動」の両方が見られたときに、どちらで評点するか、あるいは平均を取る判断をする等計測者によって異なるからであると考えられる。計測者の着目する行動を「Threat に対処した際の行動」に統一することで計測者の評点のばらつきを少なくすることの試みは、

ある程度の計測者の着眼点の一致を促したものと考えられるが、ケース 2, ケース 4 の結果から、依然としてばらつきは存在した。

その理由の一つは、計測者コメントの「Threat は起こってからずっと引きずるため、どこで評価して良いか分からない」ことに起因するものと思われる。これは、ある一つの Threat に対するクルーの行動が何回も観測される際に、計測者がどの時点のクルー行動を評価してよいかの混乱をしていることを意味している。例えば、ある Threat（例:急病人発生）が発生した際にクルーが何らかの対処をしたとしても、その Threat は継続する場合がある。その際に、クルーは最初は何らかの対処をするため CRM スキルを発揮し、その後も何らかの対処をするため CRM スキルを発揮した場合に、計測者は、どの時点の CRM スキルを計測してよいか混乱を生ずる場合がある。また別の理由として、着目する CRM スキル（「計測項目」）が異なるという問題（例：Threat X に対して、計測者 A は、「安全への主張」に 2 点、計測者 B は「安全への主張」には空欄（このスキルには着目しない）で「業務の主体的遂行」に 2 点をつける）が抽出された。

後者の問題は、計測者の着目する計測項目まで標準化する必要があるかといった課題にもつながる。

今回は主に平均値と標準偏差に基づいて分析したが、単に評点を平均するのではなく、各計測者がどの計測項目に重きを置いて評点したかといったことも考慮して分析を実施し、計測シートや標準化訓練の有効性を検証する方法もあると考えられる。もし、他の多数の計測者においても、上記 Threat X に対して着目する CRM スキルは異なっても、低い点数が評点される傾向であるなら、「Threat X に対して CRM スキルを強化すべきであり、強化すべきスキルは「安全への主張」、「業務の主体的遂行」、・・・等々の複数のスキルを強化する」といったフィードバックも可能と考えられる。今後、計測者の評点の重みを考慮した CRM スキル計測シートを検討することも考えられる。

実運用という観点では今回使用した 2 種類の計測シートは、いずれも記入するためには長い時間を要することからあまり実用的とは言えなかった。要する時間を短縮するためには、計測者の慣熟だけではなく、計測シートの簡素化が必要と思われる。しかし、計測項目数を減らすことはスキル計測の目を粗くすることになり、強化すべきスキルの抽出が困難になるという問題もある。

今後においても、計測者の着眼点がそろえられ、分析が可能で実運用にも適した計測シート開発を検討していく必要がある。

## (2) CRM スキル計測項目自体の解釈

4.6 に述べたように、ばらつきの大きい CRM スキル項目として、抽出された数が比較的多いのは「状況認識マネジメント」であった。これは、本項目の解釈が計測者によって異なる、あるいは本項目が観測しにくいことに因ると推測される。例えば、「状況認識マネジメント」の「警戒・予測」は、クルーが頭の中で警戒や予測を実施していたとしても、声に出すなど行動しない限り、計測者には計測できない。

2.5 の(1)に客観的な計測をするための工夫を述べたが、観測しにくい項目は外す、解釈が異なる項目については表現を分かりやすくするなどのより一層の工夫が必要である。

## (3) 空欄が目立つ CRM スキル計測項目

4.7 に述べたように、CRM スキル計測項目の「意見の相違の解決」や「安全への主張」において、空欄が目立つことが分かった。

この理由として、「意見の相違の解決」や「安全への主張」は、そのスキルを発揮しなければならない状況にならない限り、計測者が観測できる行動として現れにくいためと考えられる。その一方で、そのような状況にならずとも常に発揮されるべき「決定・行動のレビュー」や「チーム活動に適した雰囲気・環境作り」「タスクの配分」などにも空欄が存

在することに注意する必要がある。

## (4) チームの総合点と各 CRM スキル項目の評点との関連

図 4-6, 図 4-9 から、ケース 3 の各計測者の計測項目の評点の傾向が、2 点, 1 点が存在せず、4 点の割合が多かったことから、総合点に 4 点をつけたと考えられる。また、図 4-7, 図 4-8, 図 4-9 から、ケース 2, 4 は 1 点, 2 点を占める割合が他ケースに比べて多く、その結果として総合点に 2 点がついたと考えられる。計測者 D がケース 4 においても 2 点, 1 点の割合が増えたにも関わらず総合点は 3 点とした理由はコメントから明確であり、「2 点以下はどうしてもつけられない、総合点をつけるのは意味がない」とのことであった。2 点以下がつけられないのは、5.2(1)に述べた計測者 D のコメントから、計測者 D が査察操縦士経験があるからであると推察される。

総合点をつけることで、そのデータを蓄積し、機種別等で傾向把握を行うことも不可能ではない。しかし、傾向把握を行うなら総合点ではなくむしろ Threat や Error ごとに CRM スキル実践度合いのデータを蓄積し、分析する方が有効とも考えられる。総合点を付ける意義・フィードバック方法については、今後の検討課題である。

## 5.2 計測者の標準化

### (1) True Score 導出のための計測者同士の議論

4.1 に述べたように、計測者 4 名ともケース 2 に比較して、ケース 4 の評点のばらつきが大きくなっていた。これはケース 4 を実施する前に計測者同士での議論を行ったことが影響していると考えられる。しかし、4.3 に述べたように、4 名の計測者の標準偏差の平均を取ると、ケース 2 とケース 4 で標準偏差平均値に有意な差は認められなかった。

当初、計測者間の評点格差をなくす目的で設けた議論の場であるが、今回の実験では議論を行うことで、迷いながら 3 点としていたものが 2 点または 4 点に以降するといったように計測者各自のあいまいな部分がなくなり、計測シート等へ

の慣熟もなされ、その結果、明確な意思のもとでの評点付けが行われるようになったものと考えられる。

このことは、図 4-5 に示すケース 2, 4 の CRM スキル計測シートの集計結果、及び表 4-2 から明らかである。これらの図、表から特に計測者 A, B は、それぞれケース 2 に比べてケース 4 の方が、「3 点」を付ける割合が減少する傾向であった。査察操縦士経験のある計測者 D は、ケース 2 では「不合格を連想して 2 点以下はつけにくい」とコメントしていたが、ケース 4 において「合否の判定ではなく、再訓練という観点で 2 点以下がつけられた」とコメントしており、表 4-2 にも「1 点」の数が増えていることが分かる。これは、計測者同士の議論の中で合否の判定の観点ではなく、訓練の観点で評点するといった視点の変化に起因する。特に査察操縦士経験がある計測者が 2 点をつけにくいという結果は、文献(3)とは異なる結果であった。文献(3)では「査察操縦士経験があると、思い切って 2 点をつけられる。」という計測者も存在した。

計測者 A は、4.5 に述べたように、議論前のケース 2 では平均点をつけやすい傾向であったが、ケース 4 において 3 点の割合が減り、これも「True Score<sup>5)</sup>」導出の議論の効果であると考えられる。計測者 A のコメントから、計測者 A は、ケース 2 において、観測できない行動も標準である 3 点としていたが、議論において「観測できない行動は空欄」という考え方に統一された。

これらの考察から、今回の「True Score<sup>5)</sup>」導出のための計測者同士の議論は、主に合否の判定の観点ではなく CRM 訓練効果測定 of 観点で評点すること、観測できない行動は空欄という視点の一致、さらに平均点に偏った評点付け (Central tendency<sup>5)</sup>) の修正に影響を与えたと考えられる。

## (2) 計測者の資格要件

4.8 において、今回の結果のみでは判断できないが、当該機種の操縦経験を有する場合と、そうでない場合において評点の傾向が異なる可能性があった。

しかし、計測者 A は、4.5 に述べたよう

に、他計測者に比べて平均点の「3 点」をつける傾向があるため、当該機種の操縦経験を有するシナリオ 3 についても、他計測者が「4 点」をつける中で、単に他シナリオと同様に「3 点に」偏った評点をした可能性も考えられる。

計測者の標準化のために、2.5 の(1)に述べたように、計測者の資格要件として、「LOFT 当該機種の操縦経験が豊富であることが望ましい」と考えられる。少なくとも、4.3 に述べた計測者のコメントにも見られるように、LOFT 当該機種の SOP に関する知識は必須であると考えられる。

## (3) 今後の計測者標準化に向けて

信頼するデータを得るためには計測者の標準化訓練が必要であるが、計測者同士の観点や認識の一致をどこまで行えば標準化されたといえるのかの答えは出なかった。今回の「True Score<sup>5)</sup>」導出を目的とした計測者間の議論も 3 時間弱と、目的を達成するには短い時間であった。そのため、計測者全員の着目するスキル、評点の割合、評点結果を一致させるには不十分であったが、議論を行ったことで計測者間でのある程度の見方 (評点尺度の解釈、各点数の具体例) を揃えることができたと考えられる。言い方を変えれば、今回は各計測者が計測方法に慣熟したに過ぎず、真の標準化訓練はこれからであると思われる。LOSA の Observer 訓練<sup>4)</sup>においても、2 日～3 日と時間をかけている。

今回のように、評点付けと議論とのセットを何度か繰り返し行い、計測者間の評点差をなくしていく方法が、計測者標準化訓練として用いられている。その過程において計測者間で計測項目、評点尺度の解釈の一致が図られ、各計測者は評点で犯しやすいエラーを認識して、CRM スキル観測や評点に慣れていくと考えられる。しかしこの過程を何度繰り返すことが妥当であるかの結論は出ておらず、また、何度も繰り返したからといって評点差が完全に解消されるか疑問である。また、5.1 でも述べたように、計測者全員の着目するスキル、評点の割合までも統一する必要があるのかは今後の検討課題である。

今回の解析結果では、計測者間での評点の平均値が高い場合にばらつきが小さく、低い場合にばらつきが大きい傾向であった。これは、使用した計測シート、シナリオは違うものの、文献(3)と逆の結果であった。今後、計測者間のばらつきの要因を、多くのサンプルデータを取得して解析を実施し、抽出していく必要があると考えられる。ばらつきの要因が明確となれば、標準化訓練の参考になると考えている。

### 5.3 LOFT シナリオ開発

4.9-(2)で述べたように、Threat の種類と、それをシナリオに入れるタイミングによって、計測者の評点の傾向が一致する、または、ばらつきを生む可能性があることが分かった。

評点のばらつきの大きかったシナリオ 2 は、今回実施した模擬 LOFT の中で、テクニカル・イベントが立て続けに起こることから、プロセッサ訓練に偏り、「CRM スキル発揮が見られない」と受け取られる可能性があった。これは、テクニカルスキルが要求されるイベントは、手順が明確なためクルー間でのディスカッションをさほど要求されないことから、ある計測者は機長が独りで決定している、クルー間でプランニングしているように見られないと受け取り、ある計測者は、機長の Leadership 等に限定して評価する可能性も考えられる。この事がシナリオ 2 で、計測者間のばらつきを生んだ要因の一つであると考えられる。

受訓者が TEM のために CRM スキルを実践する事を LOFT の目的とするならば、適切な Threat の数、シナリオに挿入するタイミング及び Threat の種類や内容について注意深く検討する必要がある。計測者間でばらつきが大きい、かつどの計測者が評点しても評点の平均値が低くなるシナリオは CRM スキル訓練のシナリオとして適さないと考えられる。

## 6. まとめ

Threat and Error Management (TEM) を考慮した Crew Resource Management (CRM) 訓練の効果を測定するための CRM スキル計測手法を開発した。本手法

の妥当性を検証するため、航空会社の協力を得て模擬 Line Oriented Flight Training (LOFT) を行い、撮影した模擬 LOFT のビデオを使って、選出された 4 名の計測者による計測実験を行った。計測実験の中で試験的に「True Score<sup>5)</sup>」を出すための計測者間の議論の場を取り入れた。結論を以下にまとめる。

- ・飛行フェーズ別の計測シートと比較して Threat 別の計測シートの計測者間の標準偏差が小さくなった。この結果から、Threat 別に評点付けを行うことで、計測者間のばらつきが抑えられたと考えられる。

- ・Threat 別の CRM スキル計測シートは、ある程度の計測者の着眼点の一致を促したが、依然として、計測者間のばらつきは存在した。

- ・Threat 別の CRM スキル計測シートを用いると、飛行フェーズ別の計測シートを用いる時と比べて、計測者間で評点傾向が一致した。

- ・「True Score<sup>5)</sup>」導出のための議論の結果、迷いながら 3 点としていたものが 2 点または 4 点に以降するといったように計測者各自のあいまいな部分がなくなった。さらに、計測シート等への慣熟もなされ、その結果、明確な意思のもとでの評点付けが行われるようになった。

- ・「True Score<sup>5)</sup>」導出の議論は、主に可否の判定の観点ではなく CRM 訓練効果測定の見点で評点すること、観測できない行動は空欄とするという視点の一致、Central tendency<sup>5)</sup>の修正に影響を与えた。

- ・今回のように、評点付けと議論とのセットの過程を繰り返すことが妥当であるかどうかの結論は出なかった。

- ・Aircraft Threat が立て続けに入っているシナリオは、計測者の評点も低く、ばらつきも大きくなる可能性がある。すなわち、LOFT シナリオに組み込む Threat の種類と、それを入れるタイミングによって、計測者の評点のばらつきを生む可能性がある。

- ・計測者間での評点の平均値が高い場合は、計測者間のばらつきが小さく、低い場合は、ばらつきが大きくなる傾向であった。

以上をまとめると、計測者間のばらつ



きは、CRM スキル計測シート、計測者の標準化訓練方法（「True Score<sup>5)</sup>」導出のための計測者同士の議論の場合）、LOFT シナリオ構成に起因している可能性があることが明らかとなった。ただし、どの項目が主に関与しているかは明らかにできなかった。

今後も、計測者間のばらつきの要因を、多くのサンプルデータを取得して解析を実施し、抽出していく必要がある。

## 7. 参考文献

- 1) 飯島，野田，須藤，村岡，舩引；CRM スキル行動指標の開発：TR-1465，航空宇宙技術研究所，2003 年 7 月
- 2) 野田，飯島，津田，須藤，山森，小林；CRM スキル計測指標の開発及び評価：第 44 回飛行機シンポジウム，NO. 3G11，2005 年 10 月
- 3) H. Tsuda, T. Iijima and F. Noda, ; Development of CRM Skills Measuring Method: ICAS2006, Hamburg September, 2006.
- 4) R. L. Helmreich, et al.; The Line Operations Safety Audit (LOSA): Proc. Of 1<sup>st</sup> LOSA Week, Cathay City Hong Kong, March 2001.
- 5) D. P. Baker, C. Mulqueen and R. K. Dismukes; Training pilot instructors to assess CRM: The utility of frame-of-reference (FOR) training: International Aviation Training Symposium, Oklahoma City, September, 1999.
- 6) ヒューマンエラーを考慮した航空従事者等教育訓練方式の確立に関する調査・研究報告書，航空輸送技術研究センター，2007 年 3 月
- 7) QANTAS Flight Training, Flight Crew Observation Record FT/11
- 8) J. R. Klinect, J. A. Wilhelm and R. L. Helmreich; LOSA Error Code Book 9.0”: Proc. Of 1<sup>st</sup> LOSA Week, PP.209, Cathay City Hong Kong, March, 2001.
- 9) 野田，飯島，津田；LOFT シナリオの新設計手法について：第 45 回飛行機シンポジウム，NO. 2I4，2007 年 10 月，北九州。
- 10) Games PA, Howell JF, Pairwise multiple comparisons procedures with unequal N's

and / or variances: A Monte Carlo study. J. Educ. Statist. 1: 113-125, 1976.

## 8. 謝辞

本研究にご協力をいただいた「ヒューマンエラーを考慮した航空従事者等教育訓練方式の確立に関する調査・研究委員会」の委員をはじめワーキンググループメンバーの方々，実験に参加して頂いた被験者の皆様に深く感謝の意を表します。本研究が航空の安全に少しでも役立つよう願うものです。

## 9. 付録

本研究で使用した CRM スキル計測シートを付録 1 (Threat 別)，付録 2 (飛行フェーズ別)，CRM スキル観測シートを付録 3 に，ANOVA 結果を付録 4 に示す。

## 付録1 CRM スキル計測シート (Threat 別)

## CRM スキル計測シート(Threat 別)

LOFT 実施日 \_\_\_\_\_ 年 月 日

機種・シナリオ番号 \_\_\_\_\_

計測シート 記入日 \_\_\_\_\_ 年 月 日

計測者 氏名 \_\_\_\_\_

スレット No.	1	エラー No.	
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out <input checked="" type="checkbox"/> Takeoff / Climb <input type="checkbox"/> Cruise <input type="checkbox"/> Descend/Approach/Land	内容	MOD～SEV TURB に遭遇 (Departure Arrival Threats / 1 turbulence)

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRIFING ・客室乗務員の知識を考慮した BRIFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。(評点2以下, 4については要記入)

クルーはスレットやエラーが生じた際,

- どのように行動したか?
- その行動は適切・不適切だったか?
- 行動は効果につながったか(パフォーマンスに現れたか)?
- 行動自体が行われていなかったか?

スレット No. 2 / エラー No.			
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out <input checked="" type="checkbox"/> Takeoff / Climb <input type="checkbox"/> Cruise <input type="checkbox"/> Descend/Approach/Land	内容	AIL CONT JAM (Aircraft Threats/ 20: Aircraft malfunction)

注: ATC により Speed Control の指示を受けて INOP 状態に気づくことを想定

### CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例: システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRIFING ・客室乗務員の知識を考慮した BRIFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

スレット No. 3 / エラー No.			
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out <input checked="" type="checkbox"/> Takeoff / Climb <input type="checkbox"/> Cruise <input type="checkbox"/> Descend/Approach/Land	内容	XMC 付近の CB (Departure Threats / 1 adverse WX)

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRIFING ・客室乗務員の知識を考慮した BRIFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？



スレット No. 4 / エラー No.			
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out <input type="checkbox"/> Takeoff / Climb <input type="checkbox"/> Cruise <input checked="" type="checkbox"/> Descend/Approach/Land	内容	ATC と相反する TCAS RA (Departure Arrival Threats / 5 TCAS RA/TA)

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

スレット No. 5 / エラー No.			
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out	内 容	TCAS RA により負傷者発生 (Cabin Threats / 40 Cabin event)
	<input type="checkbox"/> Takeoff / Climb		
	<input type="checkbox"/> Cruise		
	<input checked="" type="checkbox"/> Descend/Approach/Land		

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

スレット No. 6 / エラー No.			
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out <input type="checkbox"/> Takeoff / Climb <input type="checkbox"/> Cruise <input checked="" type="checkbox"/> Descend/Approach/Land	内容	着陸飛行場選定（予測される Threat） (Operational Threats/ 32: Flight diversion)

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

スレット No.	7	エラー No.	
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out <input type="checkbox"/> Takeoff / Climb <input type="checkbox"/> Cruise <input checked="" type="checkbox"/> Descend/Approach/Land	内容	XMC 付近の CB (Arrival Threats / 1 adverse WX)

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？



スレット No. / エラー No.			
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out	内 容	(その他抽出した Error や Threats)
	<input type="checkbox"/> Takeoff / Climb		
	<input type="checkbox"/> Cruise		
	<input type="checkbox"/> Descend/Approach/Land		

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標		計測ポイント		評点			
状況認識マネジメント				1	2	3	4
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化			○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測			○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析			○	○	○	○
意思決定				1	2	3	4
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定			○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行			○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正			○	○	○	○
ワークロードマネジメント				1	2	3	4
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け			○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ			○	○	○	○
コミュニケーション				1	2	3	4
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認			○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認			○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張			○	○	○	○
チーム形成・維持				1	2	3	4
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス			○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励			○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる			○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

スレット No. / エラー No.			
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out	内 容	(その他抽出した Error や Threats)
	<input type="checkbox"/> Takeoff / Climb		
	<input type="checkbox"/> Cruise		
	<input type="checkbox"/> Descend/Approach/Land		

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評 点			
<b>状況認識マネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネジメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

スレット No. / エラー No.			
フェーズ	<input type="checkbox"/> Predeparture / Taxi-out	内 容	(その他抽出した Error や Threats)
	<input type="checkbox"/> Takeoff / Climb		
	<input type="checkbox"/> Cruise		
	<input type="checkbox"/> Descend/Approach/Land		

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標		計測ポイント		評点			
状況認識マネジメント				1	2	3	4
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化			○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測			○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析			○	○	○	○
意思決定				1	2	3	4
解決策の選択	安全運航上の最低線の確立, チームメンバの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定			○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行			○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正			○	○	○	○
ワークロードマネジメント				1	2	3	4
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け			○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ			○	○	○	○
コミュニケーション				1	2	3	4
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認			○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認			○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張			○	○	○	○
チーム形成・維持				1	2	3	4
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス			○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励			○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる			○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

## Overall Flight

全体コメント	
機長へのコメント	副操縦士へのコメント
その他・SOP の遵守状況等に関するコメント	

チームの総合点 (1, 2, 3, 4)	
-------------------------	--

## 付録2 CRM スキル計測シート(飛行フェーズ別)

**CRM スキル計測シート(飛行フェーズ別)**

LOFT 実施日 \_\_\_\_\_ 年 月 日

機種・シナリオ番号 \_\_\_\_\_

計測シート 記入日 \_\_\_\_\_ 年 月 日

計測者 氏名 \_\_\_\_\_



## Predeparture / Taxi-out

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム、コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立、チームメンバーの意見の提案・活用、選択肢の長所、短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか、自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点、意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション、「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。(評点2以下, 4については要記入)

クルーはスレットやエラーが生じた際,

- どのように行動したか?
- その行動は適切・不適切だったか?
- 行動は効果につながったか(パフォーマンスに現れたか)?
- 行動自体が行われていなかったか?

## Takeoff / Climb

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

## Cruise

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？

## Descend / Approach / Land

## CRM スキル実践度基準

1	2	3	4
望ましくない行動を観測 効果も発揮されない	望ましくない行動をやや観測 又は 望ましい行動が 観測されても効果につながらず	望ましい行動を観測 かつ 行動の効果を発揮	大変望ましい行動を観測 かつ 行動の効果を発揮

計測指標	計測ポイント	評点			
<b>状況認識マネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
状況の把握・認識の共有	個人が得た運航状況(例:システム, コックピット内外のコミュニケーション)の共有化	○	○	○	○
警戒・予測	一点集中や思い込みを避けた状況認識力の維持・状況の変化・Threat・潜在的な危険性・自分たちの陥りやすい環境の予測	○	○	○	○
問題点の分析	あらゆるリソースを活用した情報の収集・問題点の可能性の検討・客観的分析	○	○	○	○
<b>意思決定</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
解決策の選択	安全運航上の最低線の確立, チームメンバーの意見の提案・活用, 選択肢の長所, 短所の検討・目的にかなった意思決定	○	○	○	○
決定の実行	チーム全員の決定の理解・決定に沿った各自のタスクの実行	○	○	○	○
決定・行動のレビュー	決定の効果が現れているか, 自分達の行動のモニタ・修正	○	○	○	○
<b>ワークロードマネージメント</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
プランニング・優先順位付け	ワークロードが高くなる場合に備えた事前の準備・プランの確認 時間の制限とタスクの量・状況の緊急度を考慮した優先順位付け	○	○	○	○
タスクの配分	各乗員及び自動操縦装置(FMS, A/P, A/T など)へのタスクの適切な配分・クルーパフォーマンスモニタ	○	○	○	○
<b>コミュニケーション</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
情報の伝達と確認	・標準的な用語・誤解のないボディーラングエッジの使用・明瞭・適切な速さによる伝達 ・リードバック・正しく伝わっているかの確認・情報が不明確な場合の確認	○	○	○	○
ブリーフィング	・時間を取る・質問や情報提供の重要性の強調・SOP にない操作の場合の BRFING ・客室乗務員の知識を考慮した BRFING・運航の概要と相互連絡方法の確認	○	○	○	○
安全への主張	疑問点, 意見の率直な主張・危険であると感じた時の安全への主張	○	○	○	○
<b>チーム形成・維持</b>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
業務の主体的遂行	インテンションを明確に示す・リーダーの適切なサポート・アドバイス	○	○	○	○
チーム活動に適した雰囲気・環境作り	チームパフォーマンスモニタ・各乗員のワークロード確認・援助・疑問点や意見を口に出すことの奨励	○	○	○	○
意見の相違の解決	開放的なコミュニケーション, 「誰が正しいかではなく何が正しいか」に焦点を当てる	○	○	○	○

**コメント**

評点の理由について以下の観点で具体的に記述してください。（評点2以下、4については要記入）

クルーはスレットやエラーが生じた際、

- どのように行動したか？
- その行動は適切・不適切だったか？
- 行動は効果につながったか（パフォーマンスに現れたか）？
- 行動自体が行われていなかったか？



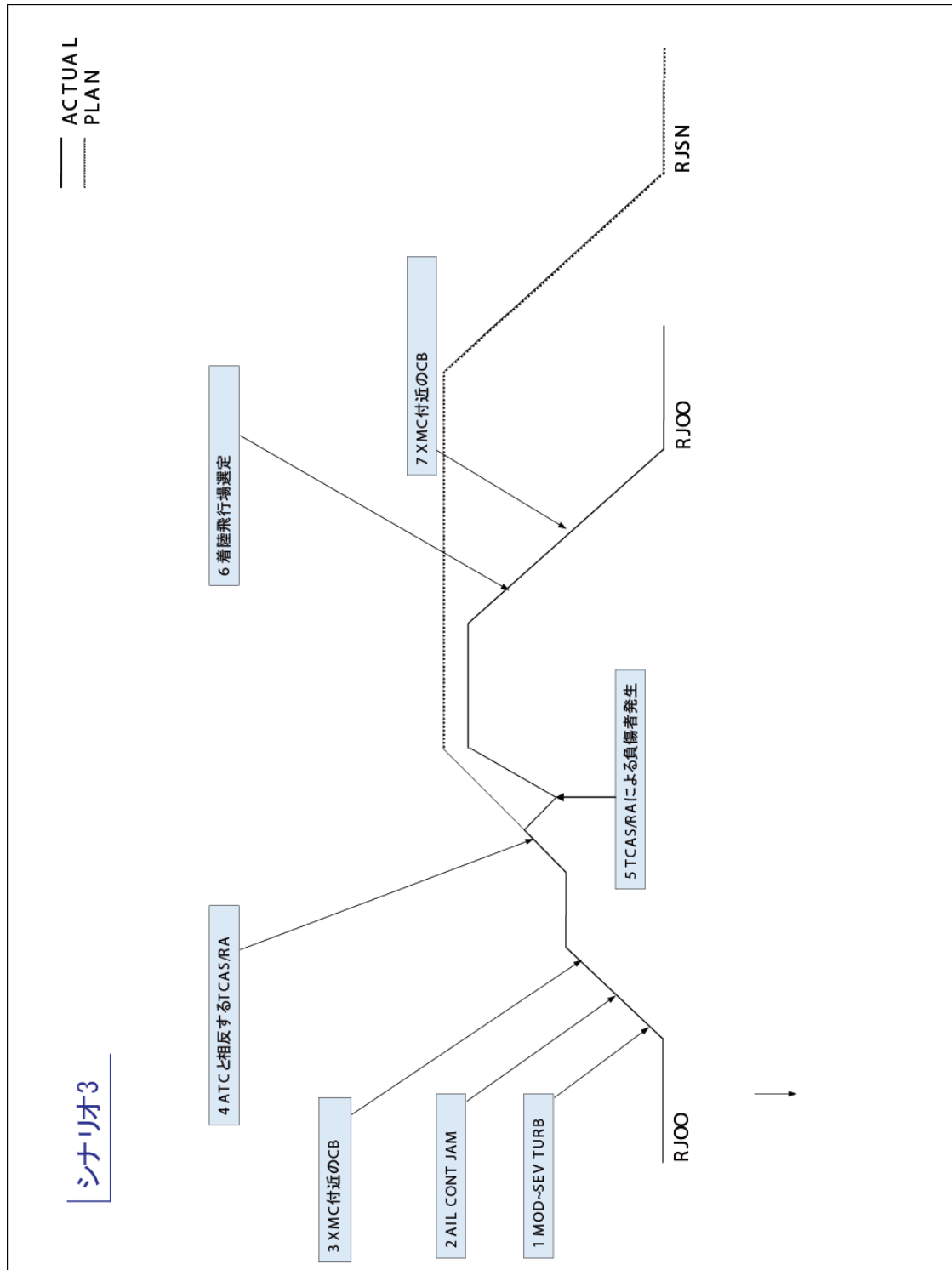
## Overall Flight

全体コメント	
機長へのコメント	副操縦士へのコメント
その他・SOP の遵守状況等に関するコメント	

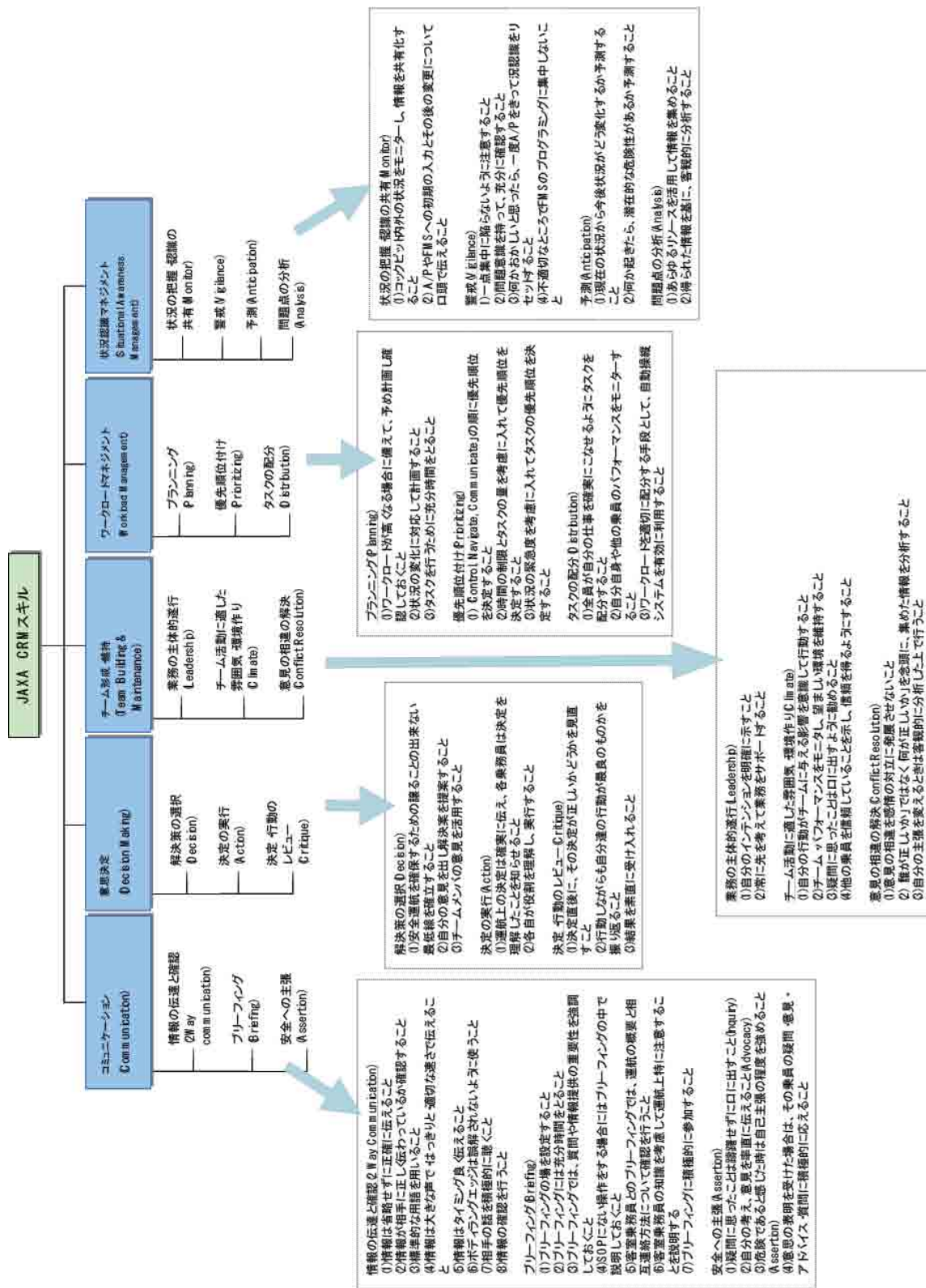
チームの総合点 (1, 2, 3, 4)	
-------------------------	--

付録3 JAXA CRM スキル観測シート(例)  
(ビデオを見ながら空いている所にメモ書きする)

(表)



(裏)



## 付録4 ANOVA 結果

## 1. 全体傾向

## 1.1 Threat 別

## (1) 計測者 A

ケース毎に有意差があるかどうかを調べるため、一元配置の ANOVA を用いてケース間の比較を行った。その結果、表 1-3 に示す F 値 ( $F=12.288$ )、有意確率 ( $P=0.000$  以下)より、計測者 A のケース毎の平均値には有意な差が存在することが分かった。

等分散性に関しては、等分散性に関する Levene 検定(表 1-2)と、平均値の等質性の検定である Welch と Brown-Forsythe(表 1-4)の検定の結果を確認した。表 1-2 に示す結果から、ケース内の等分散性を仮定した帰無仮説は、四捨五入されて 0.000 水準(0.005 以下)となっているので棄却されたことが分かる。表 1-1 から、個々の標本サイズは異なっているので、同じ標本サイズに起因する頑健性はここでは求めることができなかった。このような理由から、等分散性の過程を必要としないグループ平均値の等質性の検定である Brown-Forsythe 検定と Welch 検定を行った。表 1-4 の結果から、ケース毎の違いに高い有意な差があることを示しており、表 1-3 の結果と対応している。

表1-1 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	81	3.04	.369	.041	2.96	3.12	1	4
2	69	2.90	.349	.042	2.81	2.98	1	3
3	76	3.04	.196	.022	2.99	3.08	3	4
4	48	2.69	.512	.074	2.54	2.84	1	3
合計	274	2.94	.378	.023	2.90	2.99	1	4

表1-2 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
18.592	3	270	.000以下

※ 0.000以下は、0.000と表記

表1-3 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	4.693	3	1.564	12.288	.000
グループ内	34.373	270	.127		
合計	39.066	273			

※ 0.000以下は、0.000と表記

表1-4 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	9.088	3	125.872	.000
Brown-Forsythe	10.890	3	149.646	.000

a 漸近的 F 分布

※ 0.000以下は、0.000と表記

## 多重比較:

4 つのケースのうち、どのケースが有意に異なっているかを確認するために、多重比較を行った。多重比較には様々な手法が提案されているが、表 1-2 に示す結果から、多重比較には、データに等分散性を仮定しない手法である Games-Howell<sup>10)</sup>の手法を用いた。参考のため、他の多重比較の手法も合わせて表 1-5 に示す。他手法には、リベラルなものから一つ(LSD 手法:検定力が高いが、誤りも高くなる。t 検定を繰り返す手順であることから、多重比較法として奨励していない研究者もいる)、保守的なものから一つ(Scheffe 手法)を示した。結果については、Games-Howell<sup>10)</sup>の手法から示す。

表 1-5 の有意確率の結果から、ケース 1 とケース 4、ケース 2 とケース 3、ケース 3 とケース 4 の有意差があることが明らかとなった。

表1-5 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.138	.058	.135	-.03	.30
		3	-.002	.057	1.000	-.16	.16
		4	.350(*)	.065	.000	.17	.53
	2	1	-.138	.058	.135	-.30	.03
		3	-.141	.059	.133	-.31	.03
		4	.211(*)	.067	.021	.02	.40
	3	1	.002	.057	1.000	-.16	.16
		2	.141	.059	.133	-.03	.31
		4	.352(*)	.066	.000	.17	.54
	4	1	-.350(*)	.065	.000	-.53	-.17
		2	-.211(*)	.067	.021	-.40	-.02
		3	-.352(*)	.066	.000	-.54	-.17
LSD	1	2	.138(*)	.058	.019	.02	.25
		3	-.002	.057	.966	-.11	.11
		4	.350(*)	.065	.000	.22	.48
	2	1	-.138(*)	.058	.019	-.25	-.02
		3	-.141(*)	.059	.018	-.26	-.02
		4	.211(*)	.067	.002	.08	.34
	3	1	.002	.057	.966	-.11	.11
		2	.141(*)	.059	.018	.02	.26
		4	.352(*)	.066	.000	.22	.48
	4	1	-.350(*)	.065	.000	-.48	-.22
		2	-.211(*)	.067	.002	-.34	-.08
		3	-.352(*)	.066	.000	-.48	-.22
Games-Howell	1	2	.138	.059	.090	-.01	.29
		3	-.002	.047	1.000	-.12	.12
		4	.350(*)	.084	.001	.13	.57
	2	1	-.138	.059	.090	-.29	.01
		3	-.141(*)	.048	.020	-.27	-.02
		4	.211	.085	.071	-.01	.43
	3	1	.002	.047	1.000	-.12	.12
		2	.141(*)	.048	.020	.02	.27
		4	.352(*)	.077	.000	.15	.56
	4	1	-.350(*)	.084	.001	-.57	-.13
		2	-.211	.085	.071	-.43	.01
		3	-.352(*)	.077	.000	-.56	-.15

\* 平均の差は .05 で有意

※ 0.000以下は、0.000と表記

## (2) 計測者B

一元配置の ANOVA を実施した結果, 表 1-6 に示す F 値( $F=29.127$ ), 有意確率( $P<0.000$ )より, 計測者 B のケース毎の平均値には有意な差が存在することが分かった. 表 1-9 の結果からも, ケース毎の違いに高い有意な差があることを示していることが分かる.

表1-6 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	80	3.14	.381	.043	3.05	3.22	2	4
2	43	2.47	.667	.102	2.26	2.67	1	3
3	61	3.20	.401	.051	3.09	3.30	3	4
4	37	2.59	.599	.098	2.39	2.79	1	4
合計	221	2.93	.580	.039	2.86	3.01	1	4

表1-7 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
20.498	3	217	.000

※ 0.000以下は, 0.000と表記

表1-8 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	21.238	3	7.079	29.127	.000
グループ内	52.743	217	.243		
合計	73.982	220			

※ 0.000以下は, 0.000と表記

表1-9 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	21.957	3	93.414	.000
Brown-Forsythe	24.513	3	127.773	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

多重比較:

表 1-7 から, データに等分散性が仮定できないことから, 計測者 A と同様に, Games-Howell<sup>10)</sup>法を用いて多重比較を行った. Scheffe 法, LSD 法の結果も, 参考のため示す.

表 1-10 の Games-Howell<sup>10)</sup>手法の有意確率の結果から, ケース 1 とケース 2, ケース 4, ケース 2 とケース 3, ケース 3 とケース 4 に有意な差があることが明らかである.

表1-10 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.672(*)	.093	.000	.41	.94
		3	-.059	.084	.919	-.30	.18
		4	.543(*)	.098	.000	.27	.82
	2	1	-.672(*)	.093	.000	-.94	-.41
		3	-.732(*)	.098	.000	-1.01	-.46
		4	-.129	.111	.712	-.44	.18
	3	1	.059	.084	.919	-.18	.30
		2	.732(*)	.098	.000	.46	1.01
		4	.602(*)	.103	.000	.31	.89
	4	1	-.543(*)	.098	.000	-.82	-.27
		2	.129	.111	.712	-.18	.44
		3	-.602(*)	.103	.000	-.89	-.31
LSD	1	2	.672(*)	.093	.000	.49	.86
		3	-.059	.084	.481	-.22	.11
		4	.543(*)	.098	.000	.35	.74
	2	1	-.672(*)	.093	.000	-.86	-.49
		3	-.732(*)	.098	.000	-.93	-.54
		4	-.129	.111	.243	-.35	.09
	3	1	.059	.084	.481	-.11	.22
		2	.732(*)	.098	.000	.54	.93
		4	.602(*)	.103	.000	.40	.80
	4	1	-.543(*)	.098	.000	-.74	-.35
		2	.129	.111	.243	-.09	.35
		3	-.602(*)	.103	.000	-.80	-.40
Games-Howell	1	2	.672(*)	.110	.000	.38	.96
		3	-.059	.067	.811	-.23	.11
		4	.543(*)	.107	.000	.26	.83
	2	1	-.672(*)	.110	.000	-.96	-.38
		3	-.732(*)	.114	.000	-1.03	-.43
		4	-.129	.142	.797	-.50	.24
	3	1	.059	.067	.811	-.11	.23
		2	.732(*)	.114	.000	.43	1.03
		4	.602(*)	.111	.000	.31	.90
	4	1	-.543(*)	.107	.000	-.83	-.26
		2	.129	.142	.797	-.24	.50
		3	-.602(*)	.111	.000	-.90	-.31

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## (3) 計測者C

一元配置のANOVAを実施した結果, 表 1-13 に示す F 値( $F=45.351$ ), 有意確率( $P<0.000$ )より, 計測者 C のケース毎の平均値には有意な差が存在することが明らかとなった. 表 1-14 の結果からも, ケース毎の違いに高い有意な差があることを示していることが分かる.

表 1-11 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	62	3.05	.459	.058	2.93	3.16	2	4
2	47	2.53	.546	.080	2.37	2.69	1	3
3	62	3.53	.503	.064	3.40	3.66	3	4
4	48	2.67	.476	.069	2.53	2.80	2	3
合計	219	2.99	.628	.042	2.91	3.07	1	4

表1-12 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
17.141	3	215	.000

※ 0.000以下は, 0.000と表記

表1-13 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	33.323	3	11.108	45.351	.000
グループ内	52.659	215	.245		
合計	85.982	218			

※ 0.000以下は, 0.000と表記

表1-14 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	42.234	3	114.664	.000
Brown-Forsythe	44.811	3	199.523	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

多重比較:

表 1-12 から, データに等分散性が仮定できないことから, 計測者 A と同様に, Games-Howell<sup>10)</sup>法を用いて多重比較を行った. Scheffe 法, LSD 法の結果も, 参考のため示す.

表 1-15 の Games-Howell<sup>9)</sup>手法の有意確率の結果から, ケース 1 とケース 2, ケース 4, ケース 2 とケース 3, ケース 3 とケース 4 に有意な差があることが明らかである.



表1-15 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.672(*)	.093	.000	.41	.94
		3	-.059	.084	.919	-.30	.18
		4	.543(*)	.098	.000	.27	.82
	2	1	-.672(*)	.093	.000	-.94	-.41
		3	-.732(*)	.098	.000	-1.01	-.46
		4	-.129	.111	.712	-.44	.18
	3	1	.059	.084	.919	-.18	.30
		2	.732(*)	.098	.000	.46	1.01
		4	.602(*)	.103	.000	.31	.89
	4	1	-.543(*)	.098	.000	-.82	-.27
		2	.129	.111	.712	-.18	.44
		3	-.602(*)	.103	.000	-.89	-.31
LSD	1	2	.672(*)	.093	.000	.49	.86
		3	-.059	.084	.481	-.22	.11
		4	.543(*)	.098	.000	.35	.74
	2	1	-.672(*)	.093	.000	-.86	-.49
		3	-.732(*)	.098	.000	-.93	-.54
		4	-.129	.111	.243	-.35	.09
	3	1	.059	.084	.481	-.11	.22
		2	.732(*)	.098	.000	.54	.93
		4	.602(*)	.103	.000	.40	.80
	4	1	-.543(*)	.098	.000	-.74	-.35
		2	.129	.111	.243	-.09	.35
		3	-.602(*)	.103	.000	-.80	-.40
Games-Howell	1	2	.672(*)	.110	.000	.38	.96
		3	-.059	.067	.811	-.23	.11
		4	.543(*)	.107	.000	.26	.83
	2	1	-.672(*)	.110	.000	-.96	-.38
		3	-.732(*)	.114	.000	-1.03	-.43
		4	-.129	.142	.797	-.50	.24
	3	1	.059	.067	.811	-.11	.23
		2	.732(*)	.114	.000	.43	1.03
		4	.602(*)	.111	.000	.31	.90
	4	1	-.543(*)	.107	.000	-.83	-.26
		2	.129	.142	.797	-.24	.50
		3	-.602(*)	.111	.000	-.90	-.31

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## (4) 計測者D

一元配置のANOVAを実施した結果, 表 1-13 に示す F 値( $F=45.351$ ), 有意確率( $P=0.000$ )より, 計測者 C のケース毎の平均値には有意な差が存在することが明らかとなった. 表 1-19 の結果からも, ケース毎の違いに高い有意な差があることを示していることが分かる.

表 1-16 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	27	3.48	.509	.098	3.28	3.68	3	4
2	40	2.88	.516	.082	2.71	3.04	2	4
3	45	3.36	.484	.072	3.21	3.50	3	4
4	44	2.57	.695	.105	2.36	2.78	1	4
合計	156	3.03	.666	.053	2.93	3.14	1	4

表1-17 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
6.104	3	152	.001

表1-18 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	20.617	3	6.872	21.663	.000
グループ内	48.222	152	.317		
合計	68.840	155			

※ 0.000以下は、0.000と表記

表1-19 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	19.980	3	78.441	.000
Brown-Forsythe	22.261	3	138.202	.000

a 漸近的 F 分布

※ 0.000以下は、0.000と表記

多重比較:

表 1-17 から、データに等分散性が仮定できないことから、計測者 A と同様に、Games-Howell<sup>10)</sup>法を用いて多重比較を行った。Scheffe 法, LSD 法の結果も、参考のため示す。

表 1-20 の Games-Howell<sup>10)</sup>手法の有意確率の結果から、ケース 1 とケース 2, ケース 4, ケース 2 とケース 3, ケース 3 とケース 4 に有意な差があることが明らかである。

表1-20 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.606(*)	.140	.001	.21	1.00
		3	.126	.137	.839	-.26	.51
		4	.913(*)	.138	.000	.52	1.30
	2	1	-.606(*)	.140	.001	-1.00	-.21
		3	-.481(*)	.122	.002	-.83	-.13
		4	.307	.123	.106	-.04	.65
	3	1	-.126	.137	.839	-.51	.26
		2	.481(*)	.122	.002	.13	.83
		4	.787(*)	.119	.000	.45	1.12
	4	1	-.913(*)	.138	.000	-1.30	-.52
		2	-.307	.123	.106	-.65	.04
		3	-.787(*)	.119	.000	-1.12	-.45
LSD	1	2	.606(*)	.140	.000	.33	.88
		3	.126	.137	.360	-.14	.40
		4	.913(*)	.138	.000	.64	1.19
	2	1	-.606(*)	.140	.000	-.88	-.33
		3	-.481(*)	.122	.000	-.72	-.24
		4	.307(*)	.123	.014	.06	.55
	3	1	-.126	.137	.360	-.40	.14
		2	.481(*)	.122	.000	.24	.72
		4	.787(*)	.119	.000	.55	1.02
	4	1	-.913(*)	.138	.000	-1.19	-.64
		2	-.307(*)	.123	.014	-.55	-.06
		3	-.787(*)	.119	.000	-1.02	-.55
Games-Howell	1	2	.606(*)	.127	.000	.27	.94
		3	.126	.122	.730	-.20	.45
		4	.913(*)	.144	.000	.54	1.29
	2	1	-.606(*)	.127	.000	-.94	-.27
		3	-.481(*)	.109	.000	-.77	-.19
		4	.307	.133	.104	-.04	.66
	3	1	-.126	.122	.730	-.45	.20
		2	.481(*)	.109	.000	.19	.77
		4	.787(*)	.127	.000	.45	1.12
	4	1	-.913(*)	.144	.000	-1.29	-.54
		2	-.307	.133	.104	-.66	.04
		3	-.787(*)	.127	.000	-1.12	-.45

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## 1.2 飛行フェーズ別

Threat 別と同様の方法で, ANOVA 及び多重比較を行った. その結果を 1.2 の(1)~(4)に示す.

## (1) 計測者 A

表1-21 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	53	3.09	.295	.041	3.01	3.18	3	4
2	53	2.77	.505	.069	2.63	2.91	1	3
3	53	3.09	.295	.041	3.01	3.18	3	4
4	34	2.71	.579	.099	2.50	2.91	1	3
合計	193	2.94	.452	.033	2.87	3.00	1	4

表1-22 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
10.225	3	189	.000

※ 0.000以下は, 0.000と表記

表1-23 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	5.855	3	1.952	11.045	.000
グループ内	33.398	189	.177		
合計	39.254	192			

※ 0.000以下は, 0.000と表記

表1-24 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	9.626	3	91.462	.000
Brown-Forsythe	9.963	3	110.585	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

## 多重比較:

表1-25 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.321(*)	.082	.002	.09	.55
		3	.000	.082	1.000	-.23	.23
		4	.388(*)	.092	.001	.13	.65
	2	1	-.321(*)	.082	.002	-.55	-.09
		3	-.321(*)	.082	.002	-.55	-.09
		4	.068	.092	.910	-.19	.33
	3	1	.000	.082	1.000	-.23	.23
		2	.321(*)	.082	.002	.09	.55
		4	.388(*)	.092	.001	.13	.65
	4	1	-.388(*)	.092	.001	-.65	-.13
		2	-.068	.092	.910	-.33	.19
		3	-.388(*)	.092	.001	-.65	-.13
LSD	1	2	.321(*)	.082	.000	.16	.48
		3	.000	.082	1.000	-.16	.16
		4	.388(*)	.092	.000	.21	.57
	2	1	-.321(*)	.082	.000	-.48	-.16
		3	-.321(*)	.082	.000	-.48	-.16
		4	.068	.092	.464	-.11	.25
	3	1	.000	.082	1.000	-.16	.16
		2	.321(*)	.082	.000	.16	.48
		4	.388(*)	.092	.000	.21	.57
	4	1	-.388(*)	.092	.000	-.57	-.21
		2	-.068	.092	.464	-.25	.11
		3	-.388(*)	.092	.000	-.57	-.21
Games-Howell	1	2	.321(*)	.080	.001	.11	.53
		3	.000	.057	1.000	-.15	.15
		4	.388(*)	.107	.004	.10	.67
	2	1	-.321(*)	.080	.001	-.53	-.11
		3	-.321(*)	.080	.001	-.53	-.11
		4	.068	.121	.944	-.25	.39
	3	1	.000	.057	1.000	-.15	.15
		2	.321(*)	.080	.001	.11	.53
		4	.388(*)	.107	.004	.10	.67
	4	1	-.388(*)	.107	.004	-.67	-.10
		2	-.068	.121	.944	-.39	.25
		3	-.388(*)	.107	.004	-.67	-.10

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## (2) 計測者 B

表1-26 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	34	3.29	.462	.079	3.13	3.46	3	4
2	38	2.61	.495	.080	2.44	2.77	2	3
3	43	3.21	.412	.063	3.08	3.34	3	4
4	31	2.42	.620	.111	2.19	2.65	1	4
合計	146	2.90	.614	.051	2.80	3.00	1	4

表1-27 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
7.489	3	142	.000

※ 0.000以下は, 0.000と表記

表1-28 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	19.855	3	6.618	27.004	.000
グループ内	34.802	142	.245		
合計	54.658	145			

※ 0.000以下は, 0.000と表記

表1-29 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	24.903	3	74.299	.000
Brown-Forsythe	25.841	3	115.504	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

## 多重比較:

表1-30 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.689(*)	.117	.000	.36	1.02
		3	.085	.114	.906	-.24	.41
		4	.875(*)	.123	.000	.53	1.22
	2	1	-.689(*)	.117	.000	-1.02	-.36
		3	-.604(*)	.110	.000	-.92	-.29
		4	.186	.120	.494	-.15	.52
	3	1	-.085	.114	.906	-.41	.24
		2	.604(*)	.110	.000	.29	.92
		4	.790(*)	.117	.000	.46	1.12
	4	1	-.875(*)	.123	.000	-1.22	-.53
		2	-.186	.120	.494	-.52	.15
		3	-.790(*)	.117	.000	-1.12	-.46
LSD	1	2	.689(*)	.117	.000	.46	.92
		3	.085	.114	.457	-.14	.31
		4	.875(*)	.123	.000	.63	1.12
	2	1	-.689(*)	.117	.000	-.92	-.46
		3	-.604(*)	.110	.000	-.82	-.39
		4	.186	.120	.123	-.05	.42
	3	1	-.085	.114	.457	-.31	.14
		2	.604(*)	.110	.000	.39	.82
		4	.790(*)	.117	.000	.56	1.02
	4	1	-.875(*)	.123	.000	-1.12	-.63
		2	-.186	.120	.123	-.42	.05
		3	-.790(*)	.117	.000	-1.02	-.56
Games-Howell	1	2	.689(*)	.113	.000	.39	.99
		3	.085	.101	.836	-.18	.35
		4	.875(*)	.137	.000	.51	1.24
	2	1	-.689(*)	.113	.000	-.99	-.39
		3	-.604(*)	.102	.000	-.87	-.34
		4	.186	.137	.534	-.18	.55
	3	1	-.085	.101	.836	-.35	.18
		2	.604(*)	.102	.000	.34	.87
		4	.790(*)	.128	.000	.45	1.13
	4	1	-.875(*)	.137	.000	-1.24	-.51
		2	-.186	.137	.534	-.55	.18
		3	-.790(*)	.128	.000	-1.13	-.45

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## (3) 計測者 C

表1-31 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	38	3.11	.559	.091	2.92	3.29	2	4
2	36	2.67	.586	.098	2.47	2.86	2	4
3	52	3.81	.398	.055	3.70	3.92	3	4
4	32	2.53	.567	.100	2.33	2.74	1	3
合計	158	3.12	.735	.058	3.00	3.24	1	4

表1-32 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
5.730	3	154	.001

表1-33 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	43.091	3	14.364	53.141	.000
グループ内	41.625	154	.270		
合計	84.715	157			

※ 0.000以下は, 0.000と表記

表1-34 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	62.289	3	75.619	.000
Brown-Forsythe	49.809	3	127.455	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記



## 多重比較:

表1-35 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.439(*)	.121	.005	.10	.78
		3	-.702(*)	.111	.000	-1.02	-.39
		4	.574(*)	.125	.000	.22	.93
	2	1	-.439(*)	.121	.005	-.78	-.10
		3	-1.141(*)	.113	.000	-1.46	-.82
		4	.135	.126	.765	-.22	.49
	3	1	.702(*)	.111	.000	.39	1.02
		2	1.141(*)	.113	.000	.82	1.46
		4	1.276(*)	.117	.000	.95	1.61
	4	1	-.574(*)	.125	.000	-.93	-.22
		2	-.135	.126	.765	-.49	.22
		3	-1.276(*)	.117	.000	-1.61	-.95
LSD	1	2	.439(*)	.121	.000	.20	.68
		3	-.702(*)	.111	.000	-.92	-.48
		4	.574(*)	.125	.000	.33	.82
	2	1	-.439(*)	.121	.000	-.68	-.20
		3	-1.141(*)	.113	.000	-1.36	-.92
		4	.135	.126	.285	-.11	.38
	3	1	.702(*)	.111	.000	.48	.92
		2	1.141(*)	.113	.000	.92	1.36
		4	1.276(*)	.117	.000	1.05	1.51
	4	1	-.574(*)	.125	.000	-.82	-.33
		2	-.135	.126	.285	-.38	.11
		3	-1.276(*)	.117	.000	-1.51	-1.05
Games-Howell	1	2	.439(*)	.133	.008	.09	.79
		3	-.702(*)	.106	.000	-.98	-.42
		4	.574(*)	.135	.000	.22	.93
	2	1	-.439(*)	.133	.008	-.79	-.09
		3	-1.141(*)	.112	.000	-1.44	-.84
		4	.135	.140	.768	-.23	.50
	3	1	.702(*)	.106	.000	.42	.98
		2	1.141(*)	.112	.000	.84	1.44
		4	1.276(*)	.114	.000	.97	1.58
	4	1	-.574(*)	.135	.000	-.93	-.22
		2	-.135	.140	.768	-.50	.23
		3	-1.276(*)	.114	.000	-1.58	-.97

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## (4) 計測者 D

表1-36 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	28	3.39	.497	.094	3.20	3.59	3	4
2	33	2.67	.595	.104	2.46	2.88	2	4
3	32	3.47	.507	.090	3.29	3.65	3	4
4	31	2.42	.807	.145	2.12	2.72	1	4
合計	124	2.98	.759	.068	2.84	3.11	1	4

表1-37 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
5.625	3	120	.001

表1-38 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	25.398	3	8.466	22.314	.000
グループ内	45.529	120	.379		
合計	70.927	123			

※ 0.000以下は, 0.000と表記

表1-39 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	21.585	3	65.828	.000
Brown-Forsythe	22.460	3	101.461	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

多重比較:

表1-40 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.726(*)	.158	.000	.28	1.17
		3	-.076	.159	.973	-.53	.38
		4	.974(*)	.161	.000	.52	1.43
	2	1	-.726(*)	.158	.000	-1.17	-.28
		3	-.802(*)	.153	.000	-1.24	-.37
		4	.247	.154	.465	-.19	.68
	3	1	.076	.159	.973	-.38	.53
		2	.802(*)	.153	.000	.37	1.24
		4	1.049(*)	.155	.000	.61	1.49
	4	1	-.974(*)	.161	.000	-1.43	-.52
		2	-.247	.154	.465	-.68	.19
		3	-1.049(*)	.155	.000	-1.49	-.61
LSD	1	2	.726(*)	.158	.000	.41	1.04
		3	-.076	.159	.635	-.39	.24
		4	.974(*)	.161	.000	.66	1.29
	2	1	-.726(*)	.158	.000	-1.04	-.41
		3	-.802(*)	.153	.000	-1.10	-.50
		4	.247	.154	.111	-.06	.55
	3	1	.076	.159	.635	-.24	.39
		2	.802(*)	.153	.000	.50	1.10
		4	1.049(*)	.155	.000	.74	1.36
	4	1	-.974(*)	.161	.000	-1.29	-.66
		2	-.247	.154	.111	-.55	.06
		3	-1.049(*)	.155	.000	-1.36	-.74
Games-Howell	1	2	.726(*)	.140	.000	.36	1.10
		3	-.076	.130	.936	-.42	.27
		4	.974(*)	.173	.000	.51	1.43
	2	1	-.726(*)	.140	.000	-1.10	-.36
		3	-.802(*)	.137	.000	-1.16	-.44
		4	.247	.178	.512	-.22	.72
	3	1	.076	.130	.936	-.27	.42
		2	.802(*)	.137	.000	.44	1.16
		4	1.049(*)	.170	.000	.60	1.50
	4	1	-.974(*)	.173	.000	-1.43	-.51
		2	-.247	.178	.512	-.72	.22
		3	-1.049(*)	.170	.000	-1.50	-.60

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## 2. 提案の計測シートの有効性

各ケースについて, Threat 別シートと, 飛行フェーズ別シートにおける t 検定を実施した. 結果を表 2-1~表 2-4 に示す.

表2-1 独立サンプルの検定(ケース1)

		等分散性のための Levene の検定		2 つの母平均の差の検定					
		F 値	有意確率	t 値	自由度	有意確率 (両側)	平均値の差	差の標準誤差	差の 95% 信頼区間
									上限 下限
ケース1	等分散を仮定する.	1.702	.194	-2.105	141	.037	-.091182	.043321	-.176824 -
	等分散を仮定しない.			-2.107	112.401	.037	-.091182	.043271	-.176915 -

表2-2 独立サンプルの検定(ケース2)

		等分散性のための Levene の検定		2 つの母平均の差の検定					
		F 値	有意確率	t 値	自由度	有意確率 (両側)	平均値の差	差の標準誤差	差の 95% 信頼区間
									上限 下限
ケース2	等分散を仮定する.	.039	.844	-1.666	138	.098	-.077406	.046451	-.169255 .014442
	等分散を仮定しない.			-1.657	113.290	.100	-.077406	.046708	-.169940 .015127

表2-3 独立サンプルの検定(ケース3)

		等分散性のための Levene の検定		2 つの母平均の差の検定					
		F 値	有意確率	t 値	自由度	有意確率 (両側)	平均値の差	差の標準誤差	差の 95% 信頼区間
									上限 下限
ケース3	等分散を仮定する.	23.628	.000	-3.762	133	.000以下	-.144564	.038424	-.220565 -
	等分散を仮定しない.			-3.854	127.542	.000以下	-.144564	.037509	-.218785 -

表2-4 独立サンプルの検定(ケース4)

		等分散性のための Levene の検定		2 つの母平均の差の検定					
		F 値	有意確率	t 値	自由度	有意確率 (両側)	平均値の差	差の標準誤差	差の 95% 信頼区間
									上限 下限
ケース4	等分散を仮定する.	.514	.475	-1.592	102	.114	-.088739	.055735	-.199288 .021811
	等分散を仮定しない.			-1.576	84.984	.119	-.088739	.056310	-.200698 .023220

## 3. 計測者同士の議論の有効性

表3-1 対応サンプルの検定(Threat別)

		対応サンプルの差					t 値	自由度	有意確率 (両側)
		平均値	標準偏差	平均値の標準誤差	差の 95% 信頼区間				
					上限	下限			
ペア 1	ケース2 - ケース4	-.031774	.294330	.037380	-.106520	.042972	-.850	61	.399

表3-2 対応サンプルの検定(飛行フェーズ別)

		対応サンプルの差					t 値	自由度	有意確率 (両側)
		平均値	標準偏差	平均値の標準誤差	差の 95% 信頼区間				
					上限	下限			
ペア 1	ケース2 - ケース4	-.045952	.337013	.052002	-.150973	.059068	-.884	41	.382

## 4. 計測者の評点傾向一致の度合い

## 4-1 Threat別

表4-1 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	250	3.12	.432	.027	3.07	3.17	1	4
2	199	2.71	.544	.039	2.64	2.79	1	4
3	244	3.26	.441	.028	3.21	3.32	3	4
4	177	2.63	.570	.043	2.55	2.72	1	4
合計	870	2.97	.556	.019	2.93	3.00	1	4

表4-2 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
26.927	3	866	.000

※ 0.000以下は, 0.000と表記

表4-3 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	77.052	3	448.781	.000
Brown-Forsythe	78.559	3	726.060	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

表4-4 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	59.682	3	19.894	82.268	.000
グループ内	209.416	866	.242		
合計	269.099	869			

※ 0.000以下は, 0.000と表記

多重比較:

ANOVAを実施した結果, 有意な差が存在することが上記より確認できたため, 多重比較を行った. 等分散性が仮定できないことから, Games-Howell<sup>10)</sup>手法を用いた. その他の手法は参考のため示す.

表4-5 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) ケース	(J) ケース					上限	下限
Scheffe	1	2	.406(*)	.047	.000	.28	.54
		3	-.142(*)	.044	.016	-.27	-.02
		4	.487(*)	.048	.000	.35	.62
	2	1	-.406(*)	.047	.000	-.54	-.28
		3	-.549(*)	.047	.000	-.68	-.42
		4	.081	.051	.470	-.06	.22
	3	1	.142(*)	.044	.016	.02	.27
		2	.549(*)	.047	.000	.42	.68
		4	.630(*)	.049	.000	.49	.77
	4	1	-.487(*)	.048	.000	-.62	-.35
		2	-.081	.051	.470	-.22	.06
		3	-.630(*)	.049	.000	-.77	-.49
LSD	1	2	.406(*)	.047	.000	.31	.50
		3	-.142(*)	.044	.001	-.23	-.06
		4	.487(*)	.048	.000	.39	.58
	2	1	-.406(*)	.047	.000	-.50	-.31
		3	-.549(*)	.047	.000	-.64	-.46
		4	.081	.051	.112	-.02	.18
	3	1	.142(*)	.044	.001	.06	.23
		2	.549(*)	.047	.000	.46	.64
		4	.630(*)	.049	.000	.53	.72
	4	1	-.487(*)	.048	.000	-.58	-.39
		2	-.081	.051	.112	-.18	.02
		3	-.630(*)	.049	.000	-.72	-.53
Games-Howell	1	2	.406(*)	.047	.000	.28	.53
		3	-.142(*)	.039	.002	-.24	-.04
		4	.487(*)	.051	.000	.36	.62
	2	1	-.406(*)	.047	.000	-.53	-.28
		3	-.549(*)	.048	.000	-.67	-.43
		4	.081	.058	.499	-.07	.23
	3	1	.142(*)	.039	.002	.04	.24
		2	.549(*)	.048	.000	.43	.67
		4	.630(*)	.051	.000	.50	.76
	4	1	-.487(*)	.051	.000	-.62	-.36
		2	-.081	.058	.499	-.23	.07
		3	-.630(*)	.051	.000	-.76	-.50

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## 4-2 飛行フェーズ別

表4-6 記述統計

ケース	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
1	162	3.17	.467	.037	3.10	3.25	2	4
2	160	2.69	.540	.043	2.60	2.77	1	4
3	180	3.39	.490	.037	3.32	3.47	3	4
4	128	2.52	.652	.058	2.41	2.64	1	4
合計	630	2.98	.638	.025	2.93	3.03	1	4

表4-7 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
21.735	3	626	.000

※ 0.000以下は, 0.000と表記

表4-8 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	77.052	3	448.781	.000
Brown-Forsythe	78.559	3	726.060	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

表4-9 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	77.312	3	25.771	90.398	.000
グループ内	178.460	626	.285		
合計	255.771	629			

※ 0.000以下は, 0.000と表記

多重比較:

ANOVAを実施した結果, 有意な差が存在することが上記より確認できたため, 多重比較を行った. 等分散性が仮定できないことから, Games-Howell<sup>10)</sup>手法を用いた. その他の手法は参考のため示す.

表 4-10 多重比較

	(I) ケース	(J) ケース	平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
						上限	下限
Scheffe	1	2	.485(*)	.060	.000	.32	.65
		3	-.222(*)	.058	.002	-.38	-.06
		4	.649(*)	.063	.000	.47	.83
	2	1	-.485(*)	.060	.000	-.65	-.32
		3	-.707(*)	.058	.000	-.87	-.54
		4	.164	.063	.083	-.01	.34
	3	1	.222(*)	.058	.002	.06	.38
		2	.707(*)	.058	.000	.54	.87
		4	.871(*)	.062	.000	.70	1.04
	4	1	-.649(*)	.063	.000	-.83	-.47
		2	-.164	.063	.083	-.34	.01
		3	-.871(*)	.062	.000	-1.04	-.70
LSD	1	2	.485(*)	.060	.000	.37	.60
		3	-.222(*)	.058	.000	-.34	-.11
		4	.649(*)	.063	.000	.53	.77
	2	1	-.485(*)	.060	.000	-.60	-.37
		3	-.707(*)	.058	.000	-.82	-.59
		4	.164(*)	.063	.010	.04	.29
	3	1	.222(*)	.058	.000	.11	.34
		2	.707(*)	.058	.000	.59	.82
		4	.871(*)	.062	.000	.75	.99
	4	1	-.649(*)	.063	.000	-.77	-.53
		2	-.164(*)	.063	.010	-.29	-.04
		3	-.871(*)	.062	.000	-.99	-.75
Games-Howell	1	2	.485(*)	.056	.000	.34	.63
		3	-.222(*)	.052	.000	-.36	-.09
		4	.649(*)	.068	.000	.47	.83
	2	1	-.485(*)	.056	.000	-.63	-.34
		3	-.707(*)	.056	.000	-.85	-.56
		4	.164	.072	.104	-.02	.35
	3	1	.222(*)	.052	.000	.09	.36
		2	.707(*)	.056	.000	.56	.85
		4	.871(*)	.068	.000	.69	1.05
	4	1	-.649(*)	.068	.000	-.83	-.47
		2	-.164	.072	.104	-.35	.02
		3	-.871(*)	.068	.000	-1.05	-.69

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## 5. 各計測の評点傾向

## 5-1 Threat 別

表5-1 記述統計

計測者	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
A	274	2.94	.378	.023	2.90	2.99	1	4
B	221	2.93	.580	.039	2.86	3.01	1	4
C	219	2.99	.628	.042	2.91	3.07	1	4
D	156	3.03	.666	.053	2.93	3.14	1	4
合計	870	2.97	.556	.019	2.93	3.00	1	4



表5-2 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
14.424	3	866	.000

※ 0.000以下は, 0.000と表記

表5-3 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	1.155	3	410.361	.327
Brown-Forsythe	1.219	3	644.579	.302

a 漸近的 F 分布

表5-4 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	1.230	3	.410	1.325	.265
グループ内	267.869	866	.309		
合計	269.099	869			

## 5-2 飛行フェーズ別

表5-5 記述統計

計測者	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
A	193	2.94	.452	.033	2.87	3.00	1	4
B	146	2.90	.614	.051	2.80	3.00	1	4
C	158	3.12	.735	.058	3.00	3.24	1	4
D	124	2.98	.759	.068	2.84	3.11	1	4
合計	621	2.98	.640	.026	2.93	3.03	1	4

表5-6 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
20.430	3	617	.000

※ 0.000以下は, 0.000と表記

表5-7 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	3.058	3	305.356	.029
Brown-Forsythe	3.314	3	486.146	.020

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

表5-8 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	4.285	3	1.428	3.531	.015
グループ内	249.554	617	.404		
合計	253.839	620			

多重比較:

ANOVAを実施した結果, 有意な差が存在することが上記より確認できたため, 多重比較を行った. 等分散性が仮定できないことから, Games-Howell<sup>9)</sup>手法を用いた. その他の手法は参考のため示す.

表 5-9 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) 計測者	(J) 計測者					上限	下限
Scheffe	A	B	.034	.070	.972	-.16	.23
		C	-.182	.068	.068	-.37	.01
		D	-.038	.073	.966	-.24	.17
	B	A	-.034	.070	.972	-.23	.16
		C	-.216(*)	.073	.033	-.42	-.01
		D	-.072	.078	.837	-.29	.15
	C	A	.182	.068	.068	-.01	.37
		B	.216(*)	.073	.033	.01	.42
		D	.144	.076	.311	-.07	.36
	D	A	.038	.073	.966	-.17	.24
		B	.072	.078	.837	-.15	.29
		C	-.144	.076	.311	-.36	.07
LSD	A	B	.034	.070	.629	-.10	.17
		C	-.182(*)	.068	.008	-.32	-.05
		D	-.038	.073	.604	-.18	.11
	B	A	-.034	.070	.629	-.17	.10
		C	-.216(*)	.073	.003	-.36	-.07
		D	-.072	.078	.356	-.22	.08
	C	A	.182(*)	.068	.008	.05	.32
		B	.216(*)	.073	.003	.07	.36
		D	.144	.076	.059	-.01	.29
	D	A	.038	.073	.604	-.11	.18
		B	.072	.078	.356	-.08	.22
		C	-.144	.076	.059	-.29	.01
Games-Howell	A	B	.034	.060	.944	-.12	.19
		C	-.182(*)	.067	.034	-.36	-.01
		D	-.038	.076	.958	-.23	.16
	B	A	-.034	.060	.944	-.19	.12
		C	-.216(*)	.077	.028	-.42	-.02
		D	-.072	.085	.834	-.29	.15
	C	A	.182(*)	.067	.034	.01	.36
		B	.216(*)	.077	.028	.02	.42
		D	.144	.090	.376	-.09	.38
	D	A	.038	.076	.958	-.16	.23
		B	.072	.085	.834	-.15	.29
		C	-.144	.090	.376	-.38	.09

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## 6. 当該機種 of 操縦経験による相違

## 6-1 Threat別

表6-1 記述統計

計測者	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
A	76	3.04	.196	.022	2.99	3.08	3	4
B	61	3.20	.401	.051	3.09	3.30	3	4
C	62	3.53	.503	.064	3.40	3.66	3	4
D	45	3.36	.484	.072	3.21	3.50	3	4
合計	244	3.26	.441	.028	3.21	3.32	3	4

表6-2 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
84.967	3	240	.000

※ 0.000以下は, 0.000と表記

表6-3 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	22.145	3	106.517	.000
Brown-Forsythe	16.981	3	166.636	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

多重比較:

ANOVAを実施した結果, 有意な差が存在することが上記より確認できたため, 多重比較を行った. 等分散性が仮定できないことから, Games-Howell<sup>9)</sup>手法を用いた. その他の手法は参考のため示す.

表6-4 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) 計測者	(J) 計測者					上限	下限
Scheffe	A	B	-.157	.069	.158	-.35	.04
		C	-.493(*)	.068	.000	-.69	-.30
		D	-.316(*)	.075	.001	-.53	-.10
	B	A	.157	.069	.158	-.04	.35
		C	-.336(*)	.072	.000	-.54	-.13
		D	-.159	.078	.254	-.38	.06
	C	A	.493(*)	.068	.000	.30	.69
		B	.336(*)	.072	.000	.13	.54
		D	.177	.078	.167	-.04	.40
	D	A	.316(*)	.075	.001	.10	.53
		B	.159	.078	.254	-.06	.38
		C	-.177	.078	.167	-.40	.04
LSD	A	B	-.157(*)	.069	.023	-.29	-.02
		C	-.493(*)	.068	.000	-.63	-.36
		D	-.316(*)	.075	.000	-.46	-.17
	B	A	.157(*)	.069	.023	.02	.29
		C	-.336(*)	.072	.000	-.48	-.19
		D	-.159(*)	.078	.044	-.31	.00
	C	A	.493(*)	.068	.000	.36	.63
		B	.336(*)	.072	.000	.19	.48
		D	.177(*)	.078	.025	.02	.33
	D	A	.316(*)	.075	.000	.17	.46
		B	.159(*)	.078	.044	.00	.31
		C	-.177(*)	.078	.025	-.33	-.02
Games-Howell	A	B	-.157(*)	.056	.031	-.30	-.01
		C	-.493(*)	.068	.000	-.67	-.31
		D	-.316(*)	.076	.001	-.52	-.12
	B	A	.157(*)	.056	.031	.01	.30
		C	-.336(*)	.082	.000	-.55	-.12
		D	-.159	.089	.284	-.39	.07
	C	A	.493(*)	.068	.000	.31	.67
		B	.336(*)	.082	.000	.12	.55
		D	.177	.096	.264	-.08	.43
	D	A	.316(*)	.076	.001	.12	.52
		B	.159	.089	.284	-.07	.39
		C	-.177	.096	.264	-.43	.08

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

## 6-2 飛行フェーズ別

表6-5 記述統計

計測者	度数	平均値	標準偏差	標準誤差	平均値の 95% 信頼区間		最小値	最大値
					下限	上限		
A	53	3.09	.295	.041	3.01	3.18	3	4
B	43	3.21	.412	.063	3.08	3.34	3	4
C	52	3.81	.398	.055	3.70	3.92	3	4
D	32	3.47	.507	.090	3.29	3.65	3	4
合計	180	3.39	.490	.037	3.32	3.47	3	4

表6-6 等分散性の検定

Levene 統計量	自由度1	自由度2	有意確率
15.057	3	176	.000

※ 0.000以下は, 0.000と表記

表6-7 平均値同等性の耐久検定

	統計(a)	自由度1	自由度2	有意
Welch	37.684	3	85.413	.000
Brown-Forsythe	29.753	3	122.590	.000

a 漸近的 F 分布

※ 0.000以下は, 0.000と表記

表6-8 分散分析

	平方和	自由度	平均平方	F 値	有意確率
グループ間	15.304	3	5.101	32.425	.000
グループ内	27.690	176	.157		
合計	42.994	179			

※ 0.000以下は, 0.000と表記

多重比較:

ANOVAを実施した結果, 有意な差が存在することが上記より確認できたため, 多重比較を行った. 等分散性が仮定できないことから, Games-Howell<sup>9)</sup>手法を用いた. その他の手法は参考のため示す.

表 6-9 多重比較

			平均値の 差 (I-J)	標準誤差	有意確率	95% 信頼区間	
(I) 計測者	(J) 計測者					上限	下限
Scheffe	A	B	-.115	.081	.575	-.34	.11
		C	-.713(*)	.077	.000	-.93	-.49
		D	-.374(*)	.089	.001	-.63	-.12
	B	A	.115	.081	.575	-.11	.34
		C	-.598(*)	.082	.000	-.83	-.37
		D	-.259	.093	.053	-.52	.00
	C	A	.713(*)	.077	.000	.49	.93
		B	.598(*)	.082	.000	.37	.83
		D	.339(*)	.089	.003	.09	.59
	D	A	.374(*)	.089	.001	.12	.63
		B	.259	.093	.053	.00	.52
		C	-.339(*)	.089	.003	-.59	-.09
LSD	A	B	-.115	.081	.160	-.28	.05
		C	-.713(*)	.077	.000	-.87	-.56
		D	-.374(*)	.089	.000	-.55	-.20
	B	A	.115	.081	.160	-.05	.28
		C	-.598(*)	.082	.000	-.76	-.44
		D	-.259(*)	.093	.006	-.44	-.08
	C	A	.713(*)	.077	.000	.56	.87
		B	.598(*)	.082	.000	.44	.76
		D	.339(*)	.089	.000	.16	.51
	D	A	.374(*)	.089	.000	.20	.55
		B	.259(*)	.093	.006	.08	.44
		C	-.339(*)	.089	.000	-.51	-.16
Games-Howell	A	B	-.115	.075	.420	-.31	.08
		C	-.713(*)	.068	.000	-.89	-.53
		D	-.374(*)	.098	.002	-.64	-.11
	B	A	.115	.075	.420	-.08	.31
		C	-.598(*)	.084	.000	-.82	-.38
		D	-.259	.109	.094	-.55	.03
	C	A	.713(*)	.068	.000	.53	.89
		B	.598(*)	.084	.000	.38	.82
		D	.339(*)	.105	.011	.06	.62
	D	A	.374(*)	.098	.002	.11	.64
		B	.259	.109	.094	-.03	.55
		C	-.339(*)	.105	.011	-.62	-.06

\* 平均の差は .05 で有意

※ 0.000以下は, 0.000と表記

