17th ICCRTS

“Operationalizing C2 Agility”

Paper Title:
Framework for Assessing Coalition C4 Interoperability

Topic(s)

- ***Approaches and Organizations***
- Collaboration, Shared Awareness, and Decision Making

Author(s)

Fred Ma and Debbie Blakeney
Centre for Operational Research and Analysis
Defence R&D Canada
NDHQ, Pearkes bldg, 6 CBN
101 Colonel By Drive
Ottawa, ON
Canada K1A 0K2

Point of Contact

Fred Ma
Fred.Ma@forces.gc.ca
613-992-4550

Fred Ma and Debbie Blakeney

Framework for Assessing Coalition C4 Interoperability

***Abstract* — A command, control, computers, and communications interoperability assessment framework has been developed to track progress in interoperability over time. It covers technical, human, organizational, and policy/legal aspects of interoperability. Inputs to the framework include military and civilian personnel, domain experts, feedback from various nations, and past methodologies for characterizing interoperability. We describe the framework development and the process for analysis of interoperability.**

***Index Terms* — Interoperability, framework, Canadian Forces.**

I. Introduction

We have developed a framework for assessing the coalition interoperability of command, control, computers, and communications. A driving consideration was that significant resources are devoted to initiatives toward interoperability, and it is unclear what progress is being made on resolving the main challenges towards achieving interoperability. However, there was a strong feeling many of the limiting challenges go well beyond the technical domain, into areas such as doctrine, process compatibility, trust, culture.

This paper reports on the first iteration of an interoperability framework, which consists of identifying and organizing relevant factors and determining a rational method for assessing those factors. The proposed framework is designed to be simple so that it is transparent and facilitates consensus among the nations of a coalition can be reached. The framework is subject to change based on feedback from coalition nations after initial application. This framework development was performed by a "framework team" consisting of the authors and both civilian and military C4ISR Subject Matter Experts (SMEs) within the Department of National Defence (DND).

II. Background

The scope of the framework initially encompassed all aspects that affect C4ISR, including national capability delivery and acquisition, as well as doctrinal aspects. It was necessary to focus the scope on the following: (i) interoperability within a coalition with similar language, culture, and heritage, (ii) interoperability issues surrounding command, control, computers, and communications (C4); (iii) emphasis on joint expeditionary strategic/operational-level planning, external to theatre; and (iv) the perspective of operational personnel rather than IT/IM[1] personnel. The focus was not on the intelligence process because it was felt from discussions with military subject matter experts (SMEs) that, within the above scope, interoperability was quite mature in the intelligence domain.

From workshops with military and civilian SMEs, four overarching aspects were found to be relevant to characterizing interoperability within the scope of this study:

1. Technical aspect: Development and interconnectivity of systems;
2. Human aspect: Development of interpersonal/international relationships and trust;
3. Organizational aspect: Understanding institutional-based factors; and
4. Policy/Legal aspect: Overcoming legal and authority related factors, including policies, directives, procedures, and practices pertaining to sharing of information for situation awareness (SA) and planning at the strategic planning level.

These aspects form the basis of a hierarchical framework of interoperability factors that were identified from brainstorming sessions with SMEs, literature surveys, and feedback from relevant domestic and coalition military organizations. This paper briefly describes the development of the interoperability factors and their hierarchical relationships. The paper goes on to investigate the synthesis and assessment of these factors from lower to higher levels of the framework hierarchy.

One source of input into the working sessions was a survey of the literature on interoperability characterization [1] based, to a large degree, on works identified by Ford (Table 1). Interoperability Factors that were identified from the literature review and were within the specific scope of the study were discussed by personnel who had operational experience and helped drive the framework tree structure. Some factors were discussed more than others, but only those factors that were within the scope became a part of the framework.

Ford's surveys covered models and methodologies for measuring interoperability from the past three decades. He encountered many definitions and assessment approaches for interoperability, and noted the deep and longstanding deficiencies in interoperability, even between the U.S. Services. His surveys critiqued fourteen approaches. Ford himself favours models that permit quantitative analysis and mathematical optimization. He found that interoperability assessment approaches have had limited adoption because most

[1] Information Technology / Information Management.

Table 1: Frameworks surveyed by Ford [2][3].

	Full name	Comment
SoIM 1980	Spectrum of interoperability	Interoperability in Defense Communications, IEEE Trans. Comm., 1980 (Gilbert E. LaVean)
QoIM 1989	Quantification of Interoperability	The Quantification of Interoperability, Naval Engineers Journal, 1989 (Dennis R. Mensh, Robert S. Kite, & Paul H. Darby)
MCISI 1996	Military Communications & Information Systems Interoperability	Military Communications & Information Systems Interoperability, MILCOM, 1996 (Col. Marek Manaowicz & Col. Piotr Gajewski)
LISI 1998	Levels of Information Systems Interoperability	Levels of Information Systems Interoperability, C4ISR AWG, 1998 [C4ISR Architecture Working Group co-chaired by J6 and ASD(C3I)/CISA]
IAM 1998	Interoperability Assessment	Interoperability Assessment, Proc. 66th MORS, 1998 (revised Aug 2003) (Michael J. Leite)
OIM 1999	Organisational Interoperability Maturity Model for C2	Organisational Interoperability Maturity Model for C2, Proc. 3rd CCRTS, 1999 (Thea Clark & Richard Jones)
NMI 2003	NATO C3 Technical Architecture Reference Model for Interoperability	NATO C3 Technical Architecture Reference Model for Interoperability, 1999, 2003 [NATO Consultation, Command, and Control Agency (NC3A)]
LCI 2003	Interoperability Roadmap for C4ISR Legacy Systems	Beyond Technical Interoperability—Introducing a Reference Model for Measures of Merit for Coalition Interoperability, Proc. 8th ICCRTS,2003 (Andreas Tolk)
LCIM 2003	Levels of Conceptual Interoperability Mode	The Levels of Conceptual Interoperability Model, Proc. 2003, Fall SIW, 2003 (Andreas Tolk & James Muguira
SoSI 2004	System of Systems	System of Systems (SoSI): Final Report, CMU Tech. Report,2004 (Edwin Morris, Linda Levine, Craig Meyers, Pat Place, & Dan Plakosh)
NTI 2004	Non-technical Interoperability	Non-technical Interoperability in Multinational Forces, Proc. 9th ICCRTS, 2004 (K. Stewart, H. Clarke, P. Goillau, N. Varrall, and M.
OIAM 2005	Organisational Interoperability Agility Model	An Organisational Interoperability Agility Model, Proc. 10th ICCRTS,2004 (Gina Kingston, Suzanne Fewell, & Warren Richer)
NCW 2003	Network Centric Warfare	Incorporated into Schades's scheme (see NID)
NID 2005	NATO (C3 System) Interoperability Directive	Ford refers to Schades's 2005 conceptual delineation of levels of abstraction [4] as NID
Stop-light 2002	Stoplight	An Interoperability Roadmap for C4ISR Legacy Systems, Acq. Rev. Qtrly., 2002 (John Hamilton, Jerome Rosen, & Paul Summers)
i-Score 2007, 2008	Interoperability Score	The Interoperability Score, Proc. 2007 CSER, 2007 (Thomas Ford, John Colombi, Scott Graham, & David Jacques)

of the measurements are at relatively low (technology systems) levels, while the models dealing with high (social, organizational, procedural) levels needed further development before they are suitable for widespread use.

In the following sections, we provide an overview of our framework, briefly describe how the content and structure were developed, describe the method for generating the indicators that populate the framework, and end with conclusions and future work.

III. Framework Overview

Figure 1 shows the hierarchical structure into which the framework is organized, consisting of 3 tiers of *interoperability factors*. Tier 1 (hereafter, *T1*) consists of the four broad interoperability aspects identified above. Tier 2 (hereafter, *T2*) consists of themes within each aspect (Table 2), and represent interoperability factors of intermediate generality. Tier 3 (hereafter, *T3*) is at the bottom level of the hierarchy, and the interoperability factors at this level are the most detailed. They have been couched in the form of multiple choice questions under each T2 factor. Each question is aimed at eliciting an ordinal metric. There are 42 questions in total, examples of which are shown in Table 3. These questions form an organized questionnaire that can be fielded to coalition partners.

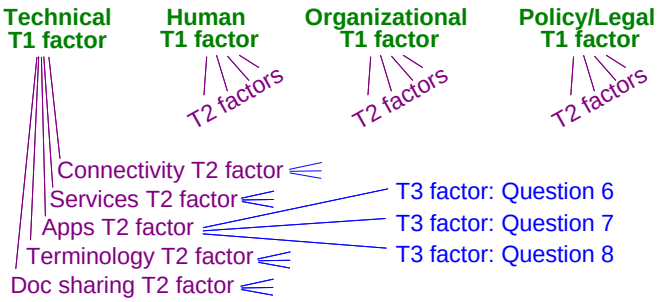

Figure 1: Hierarchical structure for the framework.

Table 2: Organization of interoperability framework at T1 and T2 levels.

	T1 Interoperability Factors			
	Technical	Human [to human relationships]	Organizational	Policy/Legal
T2 Interoperability Factors	• International connectivity • Multinational services • Applications for situational awareness and planning • Common multinational technical terminology • Multinational document sharing	• Experience with multinational cooperation • Relationships with coalition partners • Confidence in the information exchange with coalition partners	• Knowledge of coalition partners • Exchange of personnel • Accommodation of differences	• Policies • Directives • Procedures • Practices

Table 3: T3 interoperability factors and their associated metrics-generating questions (paraphrased for clarity within this report).

Technical T1 factor
• **Personnel access to coalition network:** What percentage of your nation's strategic planning staff have access to a coalition network? • **Reliability of access to coalition network:** What percentage of time are the following coalition network services for collaborative communication available? • **C2 applications interoperability:** How interoperable are the following applications in coalition operations at the strategic planning level? • **Sharing-friendliness of document defaults:** What is your national network's security default for new documents?
Human T1 factor
• **Engagement in coalition training:** In the last 12 months, has your nation's strategic planning HQ engaged in multinational collective training/exercises/operations involving coalition partners? • **Familiarity with counterparts:** How well does your nation's strategic planning staff know their coalition counterparts? • **Familiarity with coalition cultures:** Is your nation's strategic planning staff sufficiently aware of cultural differences across the coalition nations to avoid miscommunication and different understandings?
Organizational T1 factor
• **Knowledge of coalition command structures:** Is your nation's strategic planning staff familiar with the command structures of the coalition partners? • **Exchange of personnel:** In the last 12 months, did your nation have exchange/liaison officers embedded with the joint strategic planning staff of coalition partners? • **Accommodation of differences in planning processes:** Is your nation's strategic planning staff aware of the differences in coalition partners' operational planning processes?
Policy/legal T1 factor
• **Policy support of sharing:** Does your nation have policies to allow by default the sharing of information with coalition at the strategic planning level? • **Awareness of partners' operational restrictions:** Does your nation's strategic planning staff understand the different operations caveats of the coalition partners? • **Alignment of practices with sharing policies:** Does your nation's practices follow the policies, directives, and procedures to share information by default with coalition partners at the strategic planning level?

As described thus far, the framework tree structure is basically a taxonomy of interoperability factors that needs to be expanded to take input from coalition nations. Figure 2 exemplifies a coalition dashboard that is based on the vision that was articulated for the framework tasking: Stoplight colours are used to show the interoperability "health" of coalition nations under each hierarchically organized interoperability factors. In our framework, the stoplights at the bottom level of the hierarchy are based on the nations' responses to the questions (green/yellow/red indicate low/medium/high risk to interoperability). The rows correspond to the hierarchical interoperability factors while the columns correspond to nations. Most of the T2 and T3 rows are collapsed out of view, so Figure 2 shows a "drill down" into the T2 factors of one T1 factor (the Technical factor) and the questions of one T2 factor (the "Applications for Situation Awareness and Planning" factor). Note that the data shown in Figure 1 is fictitious and for demonstration purposes only.

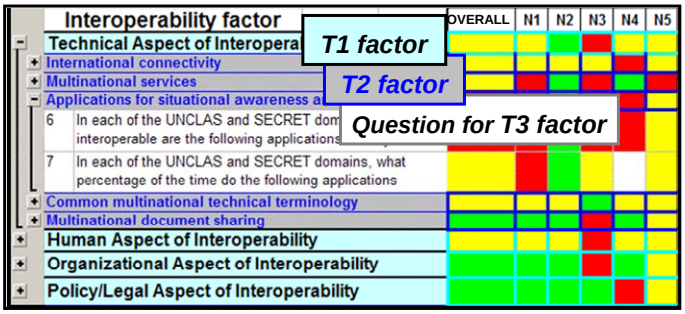

Figure 2: A coalition interoperability dashboard based on the framework hierarchy. Each "OVERALL" stoplight is a pan-coalition stoplight that is representative of all the nation-specific stoplights on the same row. The data shown is fictitious and for demonstration only.

IV. Considerations Driving Framework Content

This holistic framework requires significant input and feedback from a broad range of relevant personnel. Due to operational priorities, there are real limits on the degree to which relevant personnel can be brought to the discussions that underlie the framework. While the previous section describes the anatomy of the framework, this section describes some of the reasons why parts of the framework take the forms that they do. This includes assumptions that were needed[2], above and beyond the scoping definition, which served as a lens by which interoperability factors were selectively worked into this first iteration of framework (which will evolve in subsequent iterations based on lessons learned). It also includes the reasons for the selection of questions and the considerations behind how they are couched, as well how the T1 factors evolved.

One of the early assumptions was that the framework content would be driven by what is relevant at the "working level of the operational commands. This would be the J-staff[3] heads, (or their closest equivalents, if an HQ was not broken down by J-staff section) which we notionally took to be the Lieutenant Colonel (LCol) level.

The framework development task consisted of: (i) determining relevant T1, T2, and T3 factors; and (ii) determining a sensible way to generate metrics and associated stoplights at higher hierarchical levels from stoplights or questionnaire data at lower levels, as well as providing a way to aggregate information across the individual nations (horizontally in Figure 2) to generate overall pan-coalition interoperability indicators. As will be discussed, there can be

[2] Often by the limits to the expertise that could be obtained within the work timeframe for such a conceptually broad framework involving multiple nations.

[3] The "continental/general staff system": J1 = Personnel & Administration; J2 = Intelligence; J3 = Operations; J4 = Logistics; J5 = Plans (& Policy); J6 = Communications (& Information Systems); J7 = Doctrine & Training; J8 = Resourcing/Finance; J9 = Civil-Military Co-operation. Exact mapping of these designations may vary from organization to organization.

much subjectivity, debate, and subject matter expertise needed for part (ii), so it is expected to be more mutable than part (i). Therefore, in this paper, we will refer to part (i) as the framework and part (ii) as stoplight determination. The idea is that the stoplight determination takes data that are input at the lowest level of the framework (the questions) and populates the successively higher level tiers with interoperability indicators.

We mention above that multifaceted interoperability factors were vetted for those facets within scope. Such discussions constituted a significant part of the framework development. As an example of this vetting, consider the broad topic of technical interoperability. It could encompass openness of architectures, standardization of interfaces and data models, maturity of enterprise architecture, level of standards support by applications, and timeliness of information exchanges. Using these factors would require that we make assumptions about how they impact the degree of interoperability experienced by personnel. Instead, we attempted to maximize operational relevance by encapsulating these factors into a black box and asking about the planning person's experience of the end effects in terms of completeness and fidelity of information exchange, proper handling by the receiving application, the degree of human intervention required, and reliability (uptime). In the vetting process, timeliness was not expected to be a limiting factor at the strategic/operational level.

Focusing on the end effects is a symptomatic approach. If the dashboard brings attention to issues in a certain interoperability area, it focuses any inquiry that may be needed to identify the *causal* factors that require addressing.

One of the initial top level factors considered was the Doctrine T1 factor. Doctrine is often mentioned in the context of interoperability beyond technical systems. From discussion with the military personnel involved, it was determined that, within our scope, the relevant factors dealt with human relationships (familiarity with counterparts, familiarity with working together in operations) and familiarity with facets of partner nations that are not specific to individuals (command structure, rules that they are constrained by). The former fit well under the Human T1 factor while the latter fit under the Organizational T1 factor.

Some questions were included to confirm or dispel assumptions about how work was conducted within scope, or to provide context for the interpreting the data during analysis. For example, respondents were asked how many multinational operations and/or exercises they partook in. It is possible for operations and exercises to vary broadly in terms of scope, and different nations may refer to them by different names. Hence, respondents were asked to list the operations/exercises so that the framework team could assess how well posed the question was, whether any measures were needed to account for the diversity of operations/exercises listed, and as a basis for improving the question.

Another example of context building questions is the asking about applications for common operating pictures, targeting systems, air battle management, and C2. It was quite possible that these are not used at the strategic/operational level headquarters for the coalition of interest. The responses would either clarify this or raise questions leading to clarification. It would be better to have this information during the development of the framework and questionnaire, and feedback on a questionnaire draft can be sought from coalition nations before fielding the questionnaire. However, unless there was a prior iteration of the framework with which to demonstrate the kind of analysis that could result, one cannot assume that there would be fairly complete feedback.

Since the eventual goal is to use a relatively stable and mature version of the framework to show progress in interoperability through the years, care was taken to word the questions generically and ask about interoperability which might not yet exist. For example, questions under the Technical T1 factor (Table 3) do not refer to a specific coalition network, which was defined as any network to which all coalition nations have access. This allows the inclusion of future networks, as well as coalition enclaves of larger networks. Questions are asked about communications services, which might not exist between coalition members, but which would be very welcome.

The team also realized that it is not always constructive to ask questions that are bilateral in nature. Hence, some questions ask about the *number* of nations for which an interoperability factor applies to rather than *which* nations.

The following sections describe the quantification of T3 factors and how they aggregate (or roll up) into the higher level factors. None of the methods are mathematically or algorithmically complicated. We found that more sophisticated methods (which we briefly allude to) can account for more considerations and therefore have the potential for higher fidelity. However, they rest on more assumptions and judgement calls over which there can be greater debate. In the short term, this works against an implied purpose of framework, which is to serve as a tool of consensus how to assess the coalition state of interoperability. Therefore, reasoned, simple approaches are taken for the quantification and aggregation. This mechanisms can be upgraded on a piece-wise basis throughout the framework as the justifying operational subject matter expertise and consensus are obtained. In short, the framework should be thought of as the described hierarchy of factors framing the questions, quantification, and rollup, which can evolve as appropriate.

V. Stoplight Determination

While the nature of the T2 and T3 factors (questions) influences the choice of aggregation method (or "rollup", for brevity) to use at those levels, the interdependence between the framework and the rollup seemed most concrete at the level of the questions. Indeed, the coupling can be viewed as being tightest *within* each question; recall that the questions are multiple choice, so how the multiple choice answers to each question are defined determines the ordinal quantification of the root data to be rolled up. An example is question Q21 in the

Human T1 factor: *How well does your nation's strategic planning staff know their counterparts in the other 5-Eyes nations?* The multiple choice answers are:

A) Staff college education together
B) Worked/trained together
C) Socialized with him/her
D) Met him/her
E) Know his/her name
F) Don't know him/her
G) Not applicable (choose this for your nation only)

The answers to each question are listed in order from least-risk-to-interoperability to most-risk-to-interoperability (followed by the "Not applicable" answer). That is, from the bottom of the list to the top, the answers represent progressively decreasing risk to interoperability. For the formulation of some questions, such markers (or conditions) for demonstrating progressively decreasing levels of risk to interoperability were obvious, while for other questions (such as the one above), they were determined based on the operational expertise that could be brought into the discussions.

Without complete knowledge of what the relevant commands were like, the possible answers for some questions were defined using the best estimate of what would be meaningfully discerning or insightful.

Some questions elicit a large amount of data, and a sensible "intra-question" rollup of data is dependent on the nature of the question. For example, a question may ask the respondent nation about interoperability for a series of different communications services, each of which requires answers for every partner nation, thus yielding a 2D array of answers (Figure 3). As mentioned, the ordinality of answers (A) through (D) is based on operational expertise. The inputs to the relevant discussions, however, come from a variety of additional sources, as described in Section II[4].

Apart from questions that generate 2D tables of data, some questions generate just a series of data (a 1D vector), and some generate a single (scalar) answer. For each nation column in Figure 2, however, there is only one stoplight per question. Therefore, in order to feed T3 of the hierarchy, non-scalar data need to be condensed into 1 scalar result per question. For this reason, it is helpful to think of questions with 1D vectors of answers as having 1 additional level of depth in the hierarchy (and thus require an extra rollup step), while questions with 2D tables of answers have 2 additional levels (requiring a rollup along each dimension). For convenience, we will refer to *questions* as scalar, vector, or tabular depending on whether their *answers are scalar, vector, or tabular*.

A natural partitioning of the hierarchical rollup then occurs at the question level. The intra-question rollup *below* the question level generates one answer per question. This is quantized into a single stoplight colour (green, yellow, red) occupying one T3 cell in one nation-specific column (Figure 2). When this is done for all T3 cells, the *above* question rollup generates T1 and T2 stoplights from T2 and T3 stoplights, respectively. This partition will be referred to as the *tier rollup(s)*[5]. The following sections describe the rollup within these 2 partitions.

How interoperable are the following applications in coalition operations at the strategic planning level?

Answer for each partner nation and each application:

A) Using common applications, or as seamless
B) Exchanging information with automated assistance in conversion
C) Manual interpretation & migration of data
D) Completely not interoperable
E) Not applicable (choose this for your nation only)

Application	Nation				
	N1	N2	N3	N4	N5
Applications for common operating pictures (including	E	E	C	C	C
Targeting systems	C	E	C	C	C
Air battle management systems	C	E	C	C	B
Command and control systems	C	E	C	C	C

Figure 3: Example of a question that requires a 2D array of answers (question Q6 in the Technical T1 factor of nation *N2*'s questionnaire return).

In much of the rollup description, a common challenge is how to assign a metric of merit or worth to a result (potential or actual) that is ordinal in nature. The metric of merit may be ratio data, such as a percentage, or coarsely quantized ratio data, such as a stoplight colour. We refer this task as the *valuation* of the ordinal result, in analogy to the typical use of that word for determining the monetary value of something.

VI. Intra-question rollup

This section describes the stoplight quantization approach for questions, and the rollup for vector and tabular questions.

Whether or not a question is non-scalar (and hence requires rolling up), the final scalar result is converted into a stoplight colour. Subject matter expertise is needed to determine this mapping to a stoplight colour. Consider question Q23 in the Human T1 factor: *Is your nation's strategic planning staff sufficiently aware of cultural differences across the [coalition nations] to avoid miscommunication and different understandings?* [6] The possible answers are:

A) Communications are always properly understood

[4] The answers for the particular example in Figure 3 were motivated in part by a review of the Levels of Information Systems Interoperability surveyed in [2] and [3]. The characterization cannot be as detailed in our wholistic framework, whose scope is described in the last section as focussing on the operator's perspective.

[5] Singular or plural. We can talk about the overall aggregation scheme from tier 3 to tier 1 or individual schemes between tiers. Rollups from tier 3 to tier 2 are not the same throughout the hierarchy.

[6] This is caveated as pertaining to misunderstandings: (1) of mission intent, (2) of situational awareness info and intelligence, and (3) in cooperative planning.

B) Minimal culturally-based miscommunications
C) Occasional culturally-based miscommunications
D) Frequent culturally-based miscommunications

In an ideal framework, the relevant SMEs from the coalition nations would come to a consensus on which answers are green, yellow, and red (low, medium, and high risk to interoperability, respectively). The use of SME consensus is useful for characterizing complex, situation-dependent "things"[7] at a high level, where the applicability of more deterministic evaluation models may be limited in breadth. Typical challenges, however, are getting the relevant SMEs' input and involvement, and arriving at the consensus. Since this is a nontrivial process for which the opportunities were limited, the framework team preliminarily assigned stoplight colours to each multiple choice answer for each question, based on judgement within the team. These could be adjusted with feedback after initial application of the framework.

It was later decided that a speculative valuation of the possible answers in terms of stoplight colours was presumptuous even for a strawman. In the absence of SME consensus across a coalition, therefore, a simple mechanical method was uniformly applied to all scalar questions to generate their stoplight colours. In the context of Q23 above, this consists of: (i) mapping the possible answers (A) to (D) to equidistant points along the generic scale [0% ,100%] with the lowest answer, (D), to 0%, and the highest answer, (A), to 100% and (ii) taking the percentage figure for the actual answer provided and quantizing the number into green/yellow/red depending on whether it falls in the top, middle, or bottom third of [0% , 100%]. For step (i), the percentage points occupied by the possible answers depend on the number of possible answers, which is question specific.

It is recognized that the mapping of possible answers to equidistant points on the generic scale [0% , 100%] is nonideal in that it treats ordinal data like ratio data. Hence, the current implementation is an initial attempt for demonstration purposes to motivate the commitment of resources to fill the gap. However, it can indicate gross trends at higher levels in the framework. A SME-based consensus valuation of all multiple choice answers in terms of stoplight colours would be significant improvement. Due to the number of questions and answers, however, we acknowledge that it would also be a significant undertaking, above and beyond the typical challenges described above for SME-based consensus.

For non-scalar questions, there are extra rollup steps between steps (i) and (ii) above. In all cases, the final scalar result of the rollup is quantized into a stoplight using the same thresholds as for scalar questions. Ideally, the rollup methods are based on consensus among relevant SMEs from the coalition nations. For each non-scalar question, the framework team determined a default rollup based on (i) the knowledge available and judgement within the team and (ii) considerations of feasibility and practicality in obtaining the expertise and consensus needed for some of the rollup options. These rollups are subject to change based on feedback from relevant SMEs.

We now describe the rollup of a vector question. Consider question Q32 in the Organizational T1 factor: *Does your nation's strategic planning staff understand the different command relationships of [coalition partners]?* The possible answers are:

A) Completely
B) Partially
C) Not at all
D) Not applicable (choose this for your nation only)

The equidistant mapping of possible answers to the generic scale [0% , 100%] is the same for a scalar question. Since only (A), (B), and (C) are valid choices, they map to 100%, 50%, and 0%, respectively. However, Q32 asks for an answer for each coalition partner (Table 4). Options for rolling up vector questions are limited only by the imagination, but only 4 seemed potentially useful: (i) simple average, (ii) weighted average, (iii) lowest answer, and (iv) heuristic rules. Options (iii) and (iv) have the benefit of not requiring ratio data, so the percentages aren't even needed i.e. they can work with either the ordinal data itself or SME-based stoplight colour valuations for them.

Table 4: Intra-question rollup for a vector question Q32 in nation N2's questionnaire return. The data is fictitious.

N1	N2	N3	N4	N5
B 50%	D	C 0%	B 50%	A 100%

A simple example of a heuristic based on stoplights is "If there are 3 reds, then the rollup is red". An actual heuristic needs to be complex enough to yield an "output" result for all possible combinations of the "input" (stoplights, in this case), though it may require a sequence of steps, some of which may be arithmetic. A heuristic fragment for another question might be "If the n^{th} answer[8] in the vector is No but there are 3 Yes's, then the rollup is yellow". Heuristics are also limited only by the imagination[9]. For this initial realization of the framework, we ruled out heuristics because of the subjectivity and the SME knowledge needed in topic area of each and every vector question. However, as SME consensus is brought to bear on more and more of the rollups throughout the framework, the flexibility of heuristic rules can capture the expert knowledge. The heuristic could be as simple as recognizing an overriding combination of inputs for special evaluation, as an exception to the simpler rollups described in this paper. To rollup a vector of *n* answers, it can be as complex as an *n* dimensional table containing SME consensus valuations for each and every

[7] "Things" can literally mean anything. SME consensus is a form of the wisdom of crowds, to which any question can be submitted (though not all may be answered with equal effectiveness).

[8] The different answers in a vector might correspond to different categories of services, applications, mission types, or J-staff sections. Interoperability in some categories might be considered more critical than others, or there might be interdependencies that are difficult to represent arithmetically. Heuristic rules of arbitrary complexity can used to score the various combinations and situations.

[9] Strictly speaking, all methods are "heuristic", with the heuristic rule being defined by the method.

combination of inputs[10].

We also ruled out weighted averaging[11] because of the SME debate and consensus needed for the weights. This left averaging and lowest answer. For questions in which the vector nature arises from requiring an answer for each coalition partner, taking the rollup to be the lowest answer is suitable if coordination or cooperation between all nations is significantly impacted by any one nation's deficiency. Otherwise, the rollup was taken to be the average.

It is interesting to note that averaging is possible *because* the ordinal answers are treated as ratio data in mapping them to equidistant points in the range [0% , 100%]. This has been described above as theoretically nonideal. In principle, a more acceptable approach would be to obtain SME-based consensus on the percentage values that each ordinal answer should be worth[12], for all non-scalar questions. However, just as determining consensus stoplight colours for the possible ordinal answers to each scalar question is a sizable undertaking, determining consensus percentages for the possible ordinal answers to each non-scalar question also a sizable task. It is possible that arriving at a consensus is more challenging for the latter because there is so much more liberty with which a percentage value can be assigned compared to a stoplight valuation.

We now describe the rollup of a tabular question. Consider the question in Figure 3 from the Technical T1 factor. Answers are requested for multiple applications and multiple partner nations. Generally, the rollup for a tabular question consists of row-wise rollups to generate a column vector, followed by rollup of the column vector. We did not find any questions for which the reverse order[13] seemed more suitable. The row-wise rollups generally take the lowest answers, while the column vector rollup generally consisted of averaging. In one case (a question on whether each J-staff head knew his/her counterpart in each partner nation), the entire table was averaged. In no case was the rollup simply taken to be the lowest answer in the entire table, which would be excessively severe.

[10] This option has been used in other analyses in-house, and has the least requirement to rationalize a rollup procedure with argumentation. However, the nontrivial onus is entirely on SMEs to provide representative outputs for every possible input. For many rollups, this may not be feasible (depending on the nature of the answers). Because of the SME consensus needed, this approach is limited to small tables e.g. n=2 or 3 answers, each having 2-3 possible multiple choices.

[11] Due to the egalitarian aim of the framework, weighted averaging would not be appropriate if the scalars within the vector correspond to the coalition nations. However, weighted averaging might be appropriate in principle if (for example) a vector consisted of answers about different network services, different applications, different mission types different, J-heads, or different file types. It depends on how non-uniformly important the different column-specific things are in operations, as determined by SMEs.

[12] This (theoretical) avenue was not discussed in the context of stoplight valuations of the possible ordinal answers for scalar questions. This is because the current tier rollup uses tier 3 stoplight results (as will be described), so it is a stoplight that is the ultimate requirement for each question. If SME consensus can be obtained in assigning percentages to the possible answers for a scalar question, then it would be just as easy (if not more so) to obtain SME consensus in directly assigning stoplight colours to them.

[13] i.e. column-wise rollups followed by a rollup of the resulting row vector.

VII. Colour Averaging Across Nations at T3

Once a stoplight is generated for each question in a nation's column (Figure 2), and this is done for each nation-column, the T3 stoplights have been populated. The "Overall" column is the pan-coalition rollup for each factor, and it is taken to be the row-wise average. Taking the average rather than the lowest result implies that some mutual adjustment between nations is assumed. That is, it is assumed that nations take on roles for which their own capabilities are best suited. It is also assumed that some deficiencies in one nation can be made up for by other nation(s), at least in part.

The averaging of the stoplights is done on the green/yellow/red quantized levels rather than on the percentage values that were quantized into stoplights. The intent of this is to mimic visual method of amalgamating a tapestry of colours into a single "average" To do this averaging of stoplights, the colours green/yellow/red are first mapped the generic scale [0% , 100%] in the same manner as for the multiple choice answers to a question i.e. to 100% / 50% / 0%, respectively. The corresponding percentages for the stoplights in a dashboard row are then averaged, and the result is quantized back into a stoplight in the same manner as for a scalar question. That is, the top/middle/bottom thirds of the [0% , 100%] range map to green/yellow/red, respectively. For example, consider 5 nations, and assume that the 5 stoplights in a row for some question n are 3 reds, 1 yellow, and 1 green. The average would be red: (0%+%0+0%+50%+100%)/5 = 30%. In quantizing the average, red is [0%,33%], yellow is (33%,67%], and green is (67%,100%]. Since this approach is motivated by mimicking a human being deciding on a representative overall colour for a collection of stoplights, we refer to it as *visual/colour averaging*.

In the above example, notice that (say) the yellow input stoplight to the averaging process is mapped to 50% even though the answer to question n could have been any value in the range (33%,67%][14]. That is, we chose to colour average a set of stoplights instead of averaging the percentages that underlie the stoplights. The reasons for this is because, apart from the intent of mimicking human cognition, the use of 100% / 50% / 0% spreads the data across the full range of [0% , 100%]. This tends to ensure a certain degree of discrimination by offsetting the tendency for data distributions to increasingly concentrate around a mean with more layers of aggregation. Furthermore, colour averaging across nations at T3 makes the tier rollup more consistent throughout the framework, since colour averaging is also used in the rollup from T3 to T2, where the case for it is even stronger (discussed next section).

VIII. Tier 3-to-2 and 2-to-1 Rollups

The rollups for T3-to-T2 and T2-to-T1 occur within the individual nation columns of Figure 2.

The T3-to-T2 rollups depend on the questions that comprise each T2 factor. As in the case of intra-question rollups,

[14] After aggregation, if the question required a nonscalar answer.

however, T3-toT2 rollups consist of either averaging or taking the lowest result. Additionally, all the motivations for using colour averaging in the pan-coalition rollup apply to T3-to-T2 colour averaging. For T3-to-T2 rollups, however, there is an additional conceptual reason for preferring colour averaging: The alternative of averaging the percentages that underlie each stoplight colour presumes that the numerical values are meaningfully comparable to each other, even though they answer different questions.

The use of heuristics was described as a way to capture subject matter expertise in future intra-question rollups. The same reasoning applies to T3-to-T2 rollups. As in the case of intra-question rollups, the sophistication of a heuristic can vary broadly; more sophisticated heuristics might be limited in the number of questions that can aggregated in any one rollup due to the insight needed into interactions between questions, and due to the consensus required.

The final T2-to-T1 rollup generate T1 stoplights from their subordinate T2 stoplights. T2-to-T1 rollup deals with the biggest "chunks" of the framework. Consequently, we resort to weighted averaging, which takes into account stakeholder opinion on the relative importance of the T2 factors under each T1 factor. There are few enough T2 factors that we can reasonably expect a good response to solicitations for information about their importance.

The weights for the T2 factors could be generated via Analytic Hierarchy Process (AHP), or directly provided by questionnaire respondents. We generated the weights using a rank based approach [5] in order to avoid the pair-wise comparison of T2 factors required by AHP, and the imperfect self-consistency that can occur [6]. Rank based weights also avoid theoretical issues with directly obtaining weights from respondents – namely, questions surrounding the validity of comparing strength of preference (implied by the weights) between respondents [7].

To generate rank based weights, respondents are asked to rank the T2 factors under each T1 factor according to importance. For the T2 factors under a particular T1 factor *x* (say, the Technical T1 factor), a valuation table that is independent of the respondents is used to map each rank to a weight (this is key to avoiding respondents' strengths of preference). Using this mapping, a set of weights for the T2 factors within a T1 factor *x* can be generated for each respondent based on his/her ranking of the T2 factors. Corresponding T2 factor weights are then be averaged across respondents to yield a "consensus" set of T2 factor weights. It is these weights which are then used to average the T2 stoplights under T1 factor *x*.

There is no one absolutely "correct" valuation function for converting ranks to weights. The challenge is to come up with a reasonable one that stakeholders can agree to. In [5], the valuation function was based on the following arbitrary but intuitive conditions:

1. The weighting of "items" [15] ranked from best to worst decreases geometrically.
2. The weight of the worst item is half of what a uniform weighting would be i.e. $1/(2n)$, where n is the number of items. The rationale for this is that anything worth less is so irrelevant that it is questionable whether it would be on the list of items to rank.

Together with the requirement that the weights sum to 1, these conditions are sufficient to completely determine a generic, convex, monotonically decreasing curve whose exact shape varies with the number of items being ranked[16]. This function can be modified based on stakeholder consensus.

Superficially, a valuation function for T2 factor ranks plays a similar role to stakeholder consensus on the valuation of the ordinal answers to each question. For T2 factors, rank based weights accelerate the consensus process by providing an initial valuation function with arguably reasonable characteristics. It is possible to try a similar approach for multiple choice answers, but the applicability may be limited. For one thing, many more valuations have to be made, which hurts feasibility in the short term. As well, the questions are at a more detailed level compared to the broad enduring areas covered by the T2 factors under any one T1 factor. Therefore, the valuations may be driven by operator insight into the specific details of current operational practice rather than opinion based on broad experience. T2 factor weights fall under the latter category, which seems more amenable to the use of initial strawman valuation curves based on generic argumentation, as was done in [5]. In contrast, the questionable fit of such a general approach at the level of the multiple choice answers is just a manifestation of the tight coupling between the framework and the rollup method within the T3 questions.

To close the discussion on T2-to-T1 rollup, note that for any T1 factor, the rollup of T2 factors must be able to handle stoplights that are devoid of colour (i.e. white) due to lack of data. Simply leaving them out would have the undesired effect of equating them to 0% (red). Instead, if a white T2 factor is encountered in a rollup, its weight is adjusted to zero for that specific calculation, and the remaining T2 factor weights are scaled up by a common factor so that they add to 1.

IX. Colour Averaging Across Nations at T2 and T1

The generation of pan-coalition stoplights at T1 and T2 occur in the same way as described for T3. Figure 4 shows an example of colour averaging across nations at T3, along with a T3-to-T2 rollup. For this particular T2 factor, the intra-nation T3-to-T2 rollups within each column are done by taking the stoplight representing the highest risk to interoperability.

Note that tiers 2 and 1 have subordinate tiers, so a logical alternative to row-wise colour averaging in generating a Overall pan-coalition stoplight is to aggregate column vector of

[15] "Items" can literally be anything. In [5], the things being ranked were questions for evaluating bid proposals.

[16] The solution involves finding the zeroes of an n^{th} order polynomial, where n is the number of items being ranked.

subordinate pan-coalition stoplights. In Figure 4, this means that the stoplight with the question mark can be determined from the two stoplights beneath it rather than the stoplights on the same row. At T2 and T1, however, we still chose row-wise colour averaging for pragmatic reasons.

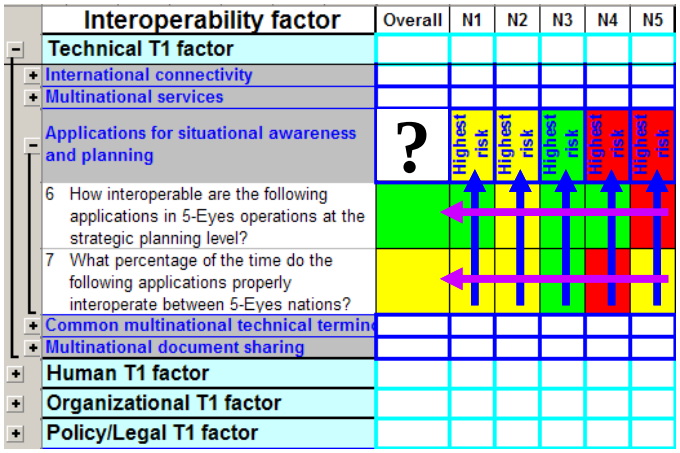

Figure 4: Colour averaging across nations. For this particular T2 factor, the vertical T3-to-T2 rollup within each column consists of taking the stoplight representing the highest risk to interoperability as opposed to taking the colour average.

The first reason is ease of transparency when presenting the analysis for discussion. For example, assume that the stakeholder discussion is focusing on a T2 stoplight. Regardless of whether a T2 factor is expanded out for visibility of the subordinate questions (Q6 and Q7 in Figure 4), stakeholders can always see how a T2 row-wise colour average was arrived at. This rationale applies at T1 as well.

The second reason apparent from Figure 4 is that if vertical aggregation is used in the Overall column, it should reflect the T3-to-T2 rollup within the nation columns. Since this is not the same for all T2 factors, this means that rollup in the Overall column will not be uniform throughout the framework, which again could complicate communication of results to stakeholders.

The final reason is based on the fact that the framework is as much a mechanism of multinational SME consensus as it is an assessment framework. It is conceivable that with further multinational SME consensus, the T3-to-T2 rollups in the nation columns could be designed based on very specific interplay that is expected between the T2 factor's questions at a national level. Depending on the exact rollup method and its rationale, it is not entirely clear that there would be the same level of agreement on the sensibility of using the same mechanism in the Overall pan-coalition column. This entire uncertainty is avoided by row-wise colour averaging.

X. Merging Multiple Returns Per Nation

Some nations may have multiple commands that fit within the focus of the interoperability framework. The multiple returns for a nation can be organized as subordinate columns in the dashboard (Figure 5). However, the framework assumes equal influence from the coalition nations when generating pan-coalition stoplights at all tiers. Therefore, multiple returns/columns from a single nation are merged so that there is only one column of stoplights per nation. At all tiers, this was done in the same manner as the row-wise colour averaging across nations in generating a pan-coalition stoplight (for the same general reasons). Within each of the multiple returns per nation, however, the tier-to-tier rollup is the same as described for a single return per nation.

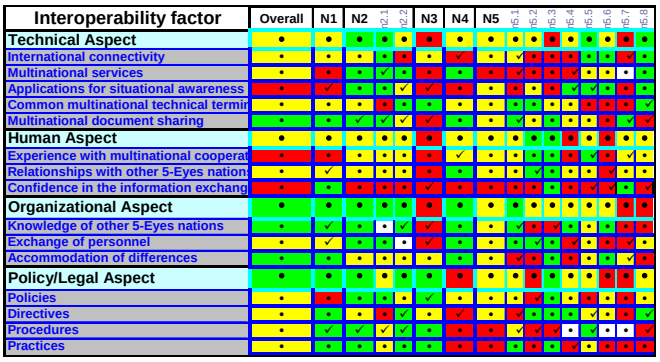

Figure 5: Hypothetically populated framework with multiple returns per nation, expanded down to the T2 level. The data is fictitious.

The dots and checkmarks Figure 5 give a crude indication of data completeness. If all answers contributing to a stoplight have been supplied, then it has a checkmark, otherwise it has a dot. Higher tier stoplights have more dots because they summarize more data, so there is a greater likelihood of one or more answers being missing.

There are some questions for which averaging across multiple returns is not appropriate. For example, some questions ask about interaction with each coalition partner. If the multiple returns correspond to regional commands, then the returns may score poorly on such a question simply because the commands largely deal only with nations that are relevant or local to their regions. The nation's forces as a whole, however, deal with all the nations named in all of the multiple returns. Since the respondent nation's aggregate forces likely have broader dealings than any single regional command, averaging the scores of the individual returns may be overly pessimistic for the nation as a whole. For such questions, it would be better to combine the responses from the multiple returns per nation before scoring and generating a stoplight.

For questions affected by this pessimism of averaging, one way to ameliorate the problem is to take the best stoplight when aggregating across multiple returns rather than the average. This is not as accurate as combining responses before scoring the question, but it may be the only solution for questions that ask only about the number of nations that a command deals with, as opposed to which nations.

In T2-to-T1 rollups, T2 factor weights are also affected by multiple returns per nation. To ensure equal influence, we want only one set of T2 factor weights per T1 factor from each nation. Therefore, corresponding T2 factor weights are averaged across the returns for a nation.

XI. The Framework as an Analytical Tool

In analogy to medical imaging or a hardware/software profiler/debugger, a hierarchical dashboard of stoplights could reveal telling patterns of interoperability health. It must be kept in mind, however, that there is no one right method for valuating and rolling up the subordinate indicators behind this visualization. The visual indicators are not absolutely "right" (or "wrong") – they are simply the result of an agreed upon methodology for characterizing a complicated state of affairs. Hence, a red stoplight does not necessarily indicate a problem, though it does warrant a drill-down into subordinate tiers (and even into the questionnaire data) to understand the cause for the red. If the red is at a high level in the framework hierarchy, however, then many subordinate factors contributed to it, which increases the likelihood that the a real gap will be found with follow-on inquiry after mining the questionnaire data. Similarly, a green does not mean that the underlying indicators can be ignored. In the absence of other considerations, however, it may make sense to look at sub-hierarchies under reds before greens.

In light of the above, the dashboard should be thought of as focusing attention on where first to drill down into the questionnaires, which contain more specific data, much of which is bilateral in nature. The stoplight indicators and the mining of the questionnaire data can guide further inquiry into *possible* interoperability gaps, which can identify causal factors and for remedial planning. As indicated, however, there is no one absolutely right way to formulate the questions that generate data for the framework, and it is possible for such follow-on inquiry to identify improvements that can be made in: (i) valuations of the questionnaire answers, (ii) rollup mechanisms, (iii) hierarchical structure, (iv) leaf node interoperability factors, and (v) the questions that generate the metrics for the factors[17]. Changes (i), (ii), and maybe (iii) can lead to further analysis iterations without the large cycle of re-fielding a questionnaire

XII. Conclusions

A first iteration of an coalition interoperability assessment framework has been developed for tracking progress in interoperability through the years. The intent was to bring under one framework the multitude of interoperability factors, including Technical, Human (to human relationships), Organizational, and a Policy/Legal aspects. For this first iteration, it was necessary to focus the scope on joint expeditionary C4 planning at the strategic/operational level, external to theatre, and from the perspective of the operational personnel. The inputs to the framework include military and civilian personnel, SMEs, feedback from various commands, and a survey of past interoperability characterization methodologies for relevant considerations.

The framework takes the form of a hierarchical organization of interoperability factors, with the most detailed factors at lowest level taking the form of metric-generating questions. These questions were packaged as a questionnaire and can be fielded to coalition nations. To get a complete picture of these factors, many of the questions are multidimensional in nature and require their own data aggregation. Some questions are designed provide insight into the operations of respondent commands, both as context for orienting the analysis and as a basis for better formulation of questions.

For maximum relevance to actual operations, many of the stoplight determination mechanisms throughout the framework would ideally make use of SME consensus on the options for data aggregation and on the valuation of nations' responses to the data gathering questions. For such a broad framework involving multiple nations, this could take some time, so we have come up with default stoplight determination mechanisms based on rational argumentation, and based on the subject matter expertise that could be obtained within the development timeframe. The framework can become more "calibrated" to real world requirements as opportunities to arise to bring subject matter expertise to bear on more and more parts of the stoplight determination (and on the framework itself).

In summary, the contributions from this work include a taxonomy of a broad range of operationally oriented interoperability factors vetted for the scope of this study, a systematic structure of rollup mechanisms which can be upgraded in a piece-wise fashion as subject matter expertise for the various parts of the framework becomes available, and a visualization tool which can guide and prioritize the descent into parts of the framework and the mining of the questionnaire data. The framework enables an iterative process in which such analytical "deep dives" generate questions about apparent interoperability gaps, rollups, or the design of the data gathering questionnaire for discussion with SMEs (domestic and coalition). Some of this inquiry may validate gaps, some may identify questions for further follow up or investigation by the coalition, and some may point to changes needed in how interoperability is characterized and rolled up in the framework.. At the end of the analysis time frame, there will a certain degree of common understanding of the state of interoperability to be captured with definitive observations and recommendations. This understanding can be illustrated by the dashboard, but with the eventual change of personnel over time, it is the observations and recommendations that endure.

XIII. Future Work

There are many ways to further the development of this framework, above and beyond the SME based calibration of the stoplight determination methods and metrics valuation. Some incremental ways are to clarify the questionnaire wording and modify the intent of some questions based on the feedback from coalition nations, either before or after initial fielding. The rollup of multiple returns per nation can be revised for the questions for which averaging is overly pessimistic.

The intent of the data completeness symbols was to give an indication of the confidence that can be ascribed to the visible

[17] As described in the final section (Future Work), the factors and the questions are not necessarily the same for later iterations for the framework.

stoplights, but it is currently very crude. A better way would be to show a partial/full moon (say, 0%, 25%, 50%, 75%, or 100%) depending on what percentage of the answers that feed a stoplight have been provided. This requires more sophisticated tracking as data is aggregated. These symbols need to be small so as not to obstruct the stoplight, so small or intricate features are to be avoided for the sake of discernibility. The level of data completeness should be communicated by a quick look at a tapestry of many such symbols, so the amount of black should be proportional to the data completeness.

A more fundamental improvement might be to flesh out the nontechnical T1 factors[18], which are currently less developed than the Technical T1 factor. In our reaching out to bring subject matter expertise into the framework development process, technical knowledge seemed to be more readily available.

The manner in which data is exploited can also be expanded. The questionnaire responses contain a wealth of structured data, and the hierarchical aggregation for this framework is but one of many ways in which they can be examined. For example, asymmetries can be identified where nation *x* reports good interoperability with nation *y* under a specific factor but the reverse is not true. Examination of these situations can determine whether they are simply the natural consequence of asymmetric relations or whether improvement can result from addressing the asymmetry. Bilateral data can also be exploited by generating social network graphs of nations or commands, where each link can be an aggregation of bilateral data from various questions[19] that are relevant to the inquiry in question. Other ways to exploit the data below tier 3 is to look for (say) services or applications that seem to be poorly scored by the majority of the nations. Finally, a variation on the current intra-question aggregation becomes obvious from the fact that a nation *x*'s questionnaire contains self-reported information about interoperability with all other partner nations, which are aggregated as a vector. It would be educational to contrast this with the vector aggregation of other nations' bilateral answers about nation *x* (as in a 360 degree feedback), and to see how this impacts the higher levels of the hierarchy.

Based on a trial fielding, a significant and immediate improvement that we will be making to the framework is to retarget the questionnaire for mass fielding. The vision for the first iteration was that there would be a lead person for each nation, whose responsibility would be to find the answers to such aggregate questions as "What percentage of your nation's strategic planning staff have access to a coalition network?" and "Is your nation's strategic planning staff familiar with the command structure of coalition partners?". Indeed, this was a necessary work scoping assumption for the timeframe of the first iteration i.e. making the question and the statement of the desired metric one and the same. However, the task of gathering the data and synthesizing the answers[20] by a nation's lead is nontrivial, which impacts the response rate, the data completeness of questionnaire returns, and the accuracy of the data. The solution is to have many operators from each command or nation complete a questionnaire and to distill the same information out of them in the analysis phase. Questions will have to be modified so that they can be immediately answered by operational staff without gathering and synthesizing information from others. Since multiple returns *per command* will be aggregated during the analysis, the aggregate information will also be more representative of the actual experience of operational staff, and less on the impressions of the a lead person (or the SMEs he/she goes to) who would otherwise come up with the aggregate information. To minimize the burden on the national leads, the questionnaire will be web based.

Even though the framework was developed to be broadly encompassing in terms of interoperability factors, it is based only on considerations of interoperability in a coalition with similar language, culture, and heritage. This framework can serve as the basis for a more expansive framework involving the nations of larger, more diverse multinational organizations. Due to the greater diversity, such a framework might put greater emphasis on differences in language, culture, command style, etc.

[18] The Technical T1 factor contains 13 of the 42 questions.

[19] Or components thereof, in the case of nonscalar questions.

[20] Or identifying and connecting with the relevant SMEs to obtain an meaningful estimates.

ACKNOWLEDGEMENTS

The authors are indebted to the rest of the framework team: LCol Pat Imai, David Yates, LCol John Blythe, Maj Samson Young. We thank Stephan Flemming for consultations on questionnaire design, Melanie Bernier for initial conception of the framework, and Ian Chapman for his review.

REFERENCES

[1] P.Bury (2011), Coalition Interoperability Measurement Frameworks Literature Survey (U), (DRDC CORA CR 2011-132), HDP Group Inc., Ottawa, ON.

[2] T.C.Ford (2007), A Survey on Interoperability Measurement, In *12th International Command and Control Research and Technology Symposium*, The Command and Control Research Program.

[3] T.C.Ford, J.M.Colombi, D.R.Jacques and S.R.Graham (2009), A General Method of Measuring Interoperability and Describing Its Impact on Operational Effectiveness, In *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 6(1), 17-32, London (UK): Sage Publications.

[4] U.Schade (2005), "Towards the Edge and Beyond: The Role of Interoperability", In *Proceedings of the 10th International Command and Control, Research and Technology Symposium*, McLean, VA, 13-16 June 2005.

[5] D.G.Hunter and E.J.Emond (2005), Analytical Support to PMO JSS: Assigning Weights to a Hierarchy Using Ordinal Rankings (U), (DRDC CORA TM 2005-32) Defence R&D Canada – CORA.

[6] J.A.Alonso and M.T.Lamata (2006), Consistency in the Analytic Hierarchy Process: A New Approach, In *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 14, No. 4 (2006) 445−459, World Scientific Publishing Company.

[7] K.J.Arrow (1951), Social Choice and Individual Values, 2nd ed., New haven & London: Yale University Press.