Defence Investment Portfolio Decisions: Insights from a National Practice Survey

John Steele
NDHQ - Carling (7S.2)
CANADA
John.Steele2@forces.gc.ca

Juha-Matti Lehtonen
National Defence University
FINLAND
juha-matti.lehtonen@mil.fi

ABSTRACT

Future uncertainty, complex military system interdependency and costly public investment to equip military forces make defence investment prioritization (DIP) among the hardest decisions any nation makes. Their difficulty and importance motivate the work of SAS-134 to survey the literature and develop guidance to help nations make DIP decisions most likely to achieve desired national outcomes. Based on the literature, we developed a 105-item survey questionnaire on national DIP practices, addressing investment planning time-frames and processes, the development of investment objectives and preferences, Operational Research (OR) methods used for analysis, the treatment of cost categories and resource constraints, and the handling of investment interactions and risk. Based on responses from 13 nations, the survey finds limited usage of OR-methods with a wide diversity in approaches. Most nations consider funding a firm constraint and some model operating budgets but no other cost categories. Diversity of DIP design indicates that procedural guidance is less useful than guiding principles, and we offer from the literature the Decision Quality construct for nations to evaluate and advance their own decision processes as need is recognized.

Keywords: Defence investment prioritization, portfolio selection, decision quality

1.0 INTRODUCTION

The research task group SAS-134 “Linking Strategic Investments & Divestments to Defence Outcomes” was initiated to identify from the literature and a survey of current international practices substantial guidance toward best practices in the prioritization of defence investments in a planned investment portfolio (PIP). Investigation of several relevant literatures informed the development and interpretation of a survey of practices in defence investment prioritization (DIP) among nations. In Section 2.0, we introduce the most relevant literatures and corresponding insights before introducing the construct within which we interpret our survey results. In Section 3.0, we describe the development and execution of the survey, and summarize the results in Sub-sections 4.1 – 4.6 before briefly concluding in Section 5.0.

2.0 INSIGHTS FROM LITERATURES RELEVANT TO DIP

We briefly summarize the DIP-relevant literature starting with Operational Research (OR), then touch on Decision Analysis (DA) before introducing and illustrating the Decision Quality construct that compactly distils DA sensibilities.

2.1 OPERATIONAL RESEARCH LITERATURE ON DEFENCE INVESTMENT

Though there are many ways to prioritize capital investments [1], the operational research practitioner readily formulates the problem as a mixed integer program to optimize a portfolio. The variations on this approach in the literature are richly varied according to the type of added complexity to be accommodated.

Brown *et al* [2] review algorithmic variations for military capital investment and Harrison *et al* [3] provide a broad recent review of the published literature.

The complexities confronting those who must prioritize defence investments are formidable, including numerous conflicting objectives, multiple stakeholders with strongly diverse perspectives and priorities, uncertain investment execution outcomes and deeply uncertain future defence needs. Hence, attaining clarity on how to plan these long-term resource allocations is not simply or even first a computational problem. The numerical solution from any calculation can only be as meaningful to a decision maker as the strength of the correspondence between the populated model and the outcome dynamics most relevant to decision-maker and stakeholder intent. While OR practitioners recognize this, the OR literature has tended to assume that the problem that calculation will address is already well defined and modelled, but getting to that point with DIP can be extremely difficult. To illustrate, the OR literature largely assumes that a portfolio must be created from scratch—that there is not already a (possibly sub-optimal) portfolio, whereas the most frequent and challenging prioritization events are reviews to reset an existing portfolio as necessary. Accordingly, it ignores the impact of investment cancellations on national industry, on national political narratives and on Defence's investment planning reputation with government. Hence, the literature, with rare exceptions[1], does not consider these less tangible losses that can accompany a gain in net modelled investment benefit through re-optimized resource allocation.

2.2 DECISION ANALYSIS

The origins of the MCDA so familiar to OR practitioners are not in OR but, as the name "multicriteria decision analysis" implies, in DA. The DA discipline focuses on strategies to manage every source of complexity that can hinder the recognition and selection of decision alternatives that best serve the decision's aims, including challenges analytical, technical, organizational and political. It explicitly addresses both the design and preparation of useful computing and the careful interpretation of results to ensure key aspects of the actual problem are not pushed aside by simplifying assumptions and the unaided limits of human reason.

The DA literature constitutes a broad and mature body of knowledge related to but distinct from that of OR in that its focus is not on algorithms and computational efficiency but on systematic approaches to understanding and characterizing uncertainty and the other sources of decision complexity—whatever has power to affect the achievement of desired decision outcomes. To be fair, the problem sets that DA and OR address are perhaps indistinguishably similar, requiring the successful practitioner in either field to know basic elements of both. The differences between them is seen in their literatures. Much OR literature presents toy problems free of complexity beyond what an algorithm will address. Case studies in OR describe decision complexities on the way to calculation for completeness' sake before addressing the algorithmic main topic. In contrast, DA emphasizes the decision complexities, themselves, and their effective management, as indispensable precursors to any meaningful calculation. Though DIP's computational challenges are real, we suggest that the lion's share of difficulty is not computational, and more specifically addressed by DA: identifying the real problem and what the decision needs to achieve, proper exploration of the feasible alternative space, realistic characterization and improvement of the available information, what can be known (and how well) about the futures toward which each alternative leads, and how to prepare for successful implementation of the best alternative.

2.3 DECISION QUALITY

In the fifty-some years since DA first emerged as a discipline to guide difficult decisions [5], its application to complex decisions of every kind has informed the development of a comprehensive framework that enables decision owners to efficiently identify and select the strongest decision alternative. Developed in the

[1] Perhaps the best example of an exception is Brown and Rosenthal (2008) [1].

1980s by SDG[2] [6], everything that can go wrong with a decision falls into one or more of the six dimensions of Decision Quality (DQ). Attaining high quality in each dimension optimizes the chances of achieving the most desired available outcomes. Hence, DQ speaks directly to "Decision Advantage in the Information Age."

Though there are sequential relationships between the six DQ dimensions, the construct is not a procedure but a set of distinct lenses through which to examine the extent of the progress made at any point toward making a high-quality decision. Understanding what high quality means in each dimension enables a useful diagnostic assessment of that aspect of any decision, affording decision owners the opportunity to improve the dimensions most needing attention before committing to a course of action. The standard for high quality is the point of diminishing returns, where additional effort to improve would not be justified by the expected results. This makes the DQ construct completely scalable, applicable both as a documented formal evaluation auditing a complex decision of high cost and impact fraught with uncertainty (such as investment prioritization), as a quick checklist for non-routine decisions requiring a little extra thought, and everything in between. Decision Quality enables decision makers to know when all the pieces are in place to make a good decision, because the construct systematically visits everything that can take a decision off the rails. As such, DQ can be profitably applied to improve every type of decision.3

The six dimensions of DQ represent aspects present in every decision:

1. The decision **Frame** defines the specific problem that the decision must address, and what a decision must accomplish. It includes the context and the most significant circumstances defining the problem. The frame separates what is out-of-scope—whether because it has already been settled by prior decisions or is to be settled later—from what is in-scope, to be settled by the current decision. It clearly specifies outcomes the decision must serve, the range of decision alternatives to be considered and the time and resources available to make and implement the decision. In short, it is the set of Ends to be served, the classes of Ways under consideration to serve them and the Means available for choosing and implementing a course of action.[4] A high quality frame ensures the decision addresses the right problem. To illustrate, consider the need for a car. The frame of the decision includes all the reasons a car is needed, satisfactory terms on which it could be obtained (renting, leasing, buying, sharing, etc.) and will be considered, the circumstances the car must or should accommodate and address (when needed, size, features, usage, image conveyed, etc.), and the resources available to acquire, maintain and operate a car (financing, parking, garage capacity, etc.).

2. Creative and feasible **Alternatives** should include several choices representing the best of the available courses of action emerging from expert and innovative survey of the possibility space, including obviously strong candidates, the best of those that emerge with creative exploration and those combining the best features of the strongest options discovered. For our example, it is the set of vehicles with strong potential to satisfy vehicular needs. This would include specific vehicles on offer, specific supply arrangements and necessary accessories (roof rack, towing package, customization, extended warranties, etc.) needed to realize the best solution.

3. **Values and trade-offs** are the specific benefits desired and relative preferences for benefits compared with each other. Values are attributes whose presence or absence in a decision alternative can be estimated—criteria reflecting the desirability of its likely outcomes. Trade-offs represent

[2] SDG (previously the Strategic Decisions Group) evolved from the Stanford Research Institute in Menlo Park, California, USA.

[3] The SAS-134 final report provides guidance on how to evaluate each DQ dimension in both generic decisions and those selecting defence investments.

[4] Lykke's much criticized formulation of strategy as Ends, Ways and Means [7], better formulates resource management.

relative outcome preferences, prioritizing competing and conflicting objectives—those features more valued in the decision outcome, those less-so and, where meaningful, how much less-so. Multiple stakeholders may bring diverging preferences to be recognized and negotiated. For a vehicle, this could deal with things like fuel consumption, seating and cargo capacity, reconfigurability for multiple uses (eg. vehicle camping); safety ratings; costs to acquire, operate and repair; ride and handling quality; acceleration at speed; available colours; etc. Values and trade-offs define what you most want and how much you are willing to sacrifice of what you want to get other things you also want. Values and trade-offs are most useful when they are quantitative. In a multi-driver household, different and conflicting properties may be desired. When they are defined in measurable terms with minimum acceptable thresholds and comparative priorities, it becomes simpler to evaluation the alternatives.

4. Relevant and reliable **Information** addresses the actual performance expected from each alternative, the resources it will require, and the degree of uncertainty in this knowledge. Tracking of information uncertainty is essential for managing outcome risks. In our example, it concerns how much of what you most want in the vehicle you should expect in each alternative. If that includes reliability, is the model only slightly refined from previous models (more predictable reliability) or freshly redesigned (of less certain reliability)? Are model reliability estimates independent or from sources owned by other automotive interests? Has a test-drive confirmed satisfactory ergonomics? Did you bring strong contenders to your mechanic for evaluation? (How much confidence do you have in your mechanic?) Will the driving be of an unusual nature, undermining "average" driving assumptions? Is the maker going out of business or dropping the model from its line-up (affecting availability and costs of parts)? How often is inclement weather likely to prevent your use of the vehicle when needed?

5. **Sound Reasoning** is about analysis to discover the most preferred alternatives—using what is known (information) about available choices (alternatives) to realize what is desired (values). It acknowledges information uncertainty and value uncertainty from stakeholder disagreement, along with any simplifying assumptions embedded in the analysis. Reasoning must judge each alternative against the requirements in the decision frame. For our example, reasoning judges how well the combinations of features in the strongest alternatives are expected to turn out for as long as the vehicle is driven. Will more than one driver want the car at the same time? (Perhaps a multi-use vehicle will not be an advantage.) Do all the alternatives meet minimum acceptable thresholds? Does the cargo capacity come with a boxy look that spoils its use on evenings out? Which information on each of the alternatives is most likely to be off, and how much is that likely to matter? What would not buying the most expensive alternative let you do with the money not spent?

Commitment to action is what implements the decision. It comes from shared stakeholder and implementer understanding and agreement that the decision is right. Otherwise, implementation can miss the target. A review of attainment on the other five DQ dimensions, acknowledging and leveraging key stakeholder insights generates implementation buy-in. It culminates in a realistic plan that sets up decision success. For our example, have you involved all the drivers in the decision process or presumed you know what they needed? Have their questions all been asked and answered to everyone's satisfaction? Does having judged between the strongest alternatives make you want to revise a couple of value trade-offs? Has what you said you needed in a vehicle changed much from what you said at the start, and is the preferred choice the best solution to the currently understood problem? Does the preferred choice require modifications to parking, the garage, property line clarifications, negotiation with neighbours affected by the size or noise of the vehicle? Is financing in place? If one driver made a personal behaviour commitment to participate in the selection process, are consequences of non-performance specified and agreed? Are you satisfied that you know what to expect with the choice and are ready for the most foreseeable surprises?

2.4 PLANNED INVESTMENT PORTFOLIO DECISION QUALITY

Decisions that adjust defence capital investment allocations present specific challenges along each of the same six dimension:

1. The **Frame** must recognize the nation's interests, external and internal, as a political, economic, geographic and cultural entity with strategic relationships, and all the implied tasks to defend and advance those interests in an uncertain future. It includes estimated current and future resource assumptions as well as capability and other investment objectives over time.

2. The **Alternatives** to be considered are not investments but investment combinations that must, with acceptable risk, affordably meet national needs for defence-relevant materiel.

3. The **Values** that investment combinations must serve are sufficiently complex and conflicting to confound defining a comprehensive and robust objective function. This confines standard portfolio optimization to a supporting role. Investment deliverable benefits interact strongly with other capability inputs and legacy systems per operational context.

4. Portfolio **Information**: besides uncertain project management outcomes (performance traded away to reduce costs and delays, aborted investments to regain affordability), future operational demands are deeply uncertain, with short and incomplete warning of true capability needs.

5. **Sound Reasoning** cannot escape some risk acceptance on every front, settling for adequate service of the most important objectives, accepting opportunity costs and risks that are not so much optimal as the least intolerable, no worse than absolutely necessary.

6. **Commitment** to the selected portfolio is wholly determined by process: no stronger than the perceived quality of its preparation for an uncertain future, no broader than the stakeholder perspectives it included and no more consequential than the embodiment of success factor insights in the plan it produces.

3.0 SURVEY DESIGN AND ADMINISTRATION

The research task group SAS-134 designed and conducted a survey exploring how nations prioritize their defence strategic investments. Key influences on survey design were the work of Tate and Thompson [8][9]. The aim was to learn from international experience best practices in dealing with the complexity that always attends major defence investment decisions. The topics addressed in the survey included: investment planning time-frame and process, how preferences and end objectives are considered, which OR-methods are facilitating the portfolio decision, what cost categories and resource constraints are considered, and treatment of complex interactions and risks.

The survey was designed as an expert survey, which assumes that there is one specific approach that a nation applies in defence investment prioritization and that only those involved and therefore knowledgeable can be surveyed. Hence, only one response is solicited from each nation. Accordingly, the reliability of the survey is evaluated not by the spread of responses but by the qualifications of the responding experts. Respondents were asked to indicate their role and how recently involved in the process. The survey sought respondents involved in the past five years in making or providing analytical support to cost-informed decisions prioritizing defence capital investments that will deliver, among other things, major capital equipment. As the questions covered a broad range of topics, the instructions recommended that the survey be completed by a small team with experience in both investment prioritization decision processes and the supporting analyses.

In September 2019 a pilot survey was administered twice to assess clarity of the questions, respondent knowledge requirements and completion time. Based on the results some changes were made for the final survey. The final survey included 105 questions in English. Its response menus, depending on the question, were largely either five-item Likert-type levels of agreement or disagreement, other constructed levels of analysis or nominal yes/no answers. Each question also included a free text field inviting related comments. Since a good knowledge of English does not necessarily imply familiarity with English decision analytic terminology, a glossary was included for clarity of technical terms that might only be familiar in the respondent's native language.

The survey was delivered as a MicroSoft® Excel™ workbook with conditional formatting to indicate missing responses, although some respondents preferred to fill in an accompanying Portable Document Format file which lacked these cues. The professional networks of SAS-134 members and more than one round of assistance from the SAS Panel provided initial contacts reached through email. A formal follow-up letter was sent from NATO Science & Technology Organization to help obtain national administrative approvals to complete the survey, though at least one large nation declined on the basis of the survey's breadth and complexity, and the infeasibility of assembling sufficient expertise. Between February 2019 and June 2020, surveys were received from 13 nations in NATO, the Partnership-for-Peace program and Enhanced Opportunity Partner nations, although it did not include every nation participating in SAS-134. With minor exceptions, all surveys were complete. All respondents indicated their role in previous DIP and most indicated the number of years since their involvement in each role with all but one of those involved in the previous 5 years. Reading through survey responses, one (maybe two) surveys with repetitive responses in some sections and few or no comments may have come from unfortunate staff officers doing their best. In total, 433 comments were received, including at least some from every nation, enhancing response interpretation and pointing out questions that were, for some countries, not easy to answer or meaningful.

With no previous international conversation specific to defence investment prioritization beyond Capability Based Planning, which has sometimes excluded those deciding investment priority, national reluctance to disclose prioritization practices was anticipated. This was addressed by promising all respondents that all resulting analyses would conceal the identities of respondents and their nations. However, if we define a large nation threshold at having more than 20 million inhabitants and a gross domestic product of at least 10^{12} US dollars, then six responding nations were large and seven small. Eight survey respondents represented ministries of defence, three defence forces and two research organizations supporting defence decision-making. To some extent, this may reflect differences of expertise and process familiarity across the different topics surveyed.

4.0 SURVEY RESULTS

The structure of the survey instrument was topical, having been designed and administration begun before the cogency of Decision Quality to the investment portfolio problem was fully appreciated. More detailed survey results and the survey instrument are reported in Annexes C and D of the SAS-134 final report, respectively. The following are the most salient survey results.

4.1 Decision Frame

Responding nations indicated that their PIPs spanned periods that ranged from just three to twenty years, divided into anywhere from one to five sub-periods. The reported time between PIP reviews ranged from annually, to every six years. Countries mostly adopt rolling multi-year plans which are revised each time a PIP review is conducted.

The most commonly described temporal for national PIPs involved three PIP sub-periods, with the first sub-period building the following year's defence budget and giving detailed costing for the next fiscal year with

following multi-year funding forecasts. The second PIP sub-period includes Preparation (2-4 year time horizons) where cost estimates are less certain, combined with Acquisition Planning, and any necessary revisions of Defense Plans. The third sub-period emphasizes the most expensive and risky Investments, Research and Development (R&D) and Acquisition, and any necessary revisions of Defense Plans. As there were large differences with respect to PIP time spans, this must bear on the kinds of issues included in a nation's planning process. For example, if a PIP stretches only a few years into the future, proposed investments under consideration are likely closer to implementation and therefore lend themselves to better cost estimates than may be the case for planned investments in the distant future (e.g. 10 years or more).

Figure 1 shows that also other items alongside major equipment acquisitions from the investment prioritization process. Corresponding to the concept of investment in a "Total System," roughly half of the nations included force size decisions in their DIP process. Other included initiatives ranged from software acquisitions (12 nations), land acquisitions (8 nations) and divestments (7 nations).

Figure 2 gives indications of some decision framing behaviours by surveyed nations. The green bars to the right in Figure 2 indicate the number of countries that tended to agree or strongly agreed that statements applied to their nation, while the orange and red bars to the left indicate nations that tended to disagree or strongly disagreed with the statement, respectively. Yellow response bars are centrally aligned indicating neither agreement nor disagreement. Ten countries indicate they solicit preferences and priorities from PIP decision makers and participating stakeholders, and eight nations agree they develop preferences into criteria. Eight indicated that criteria are stated as end-state objectives, and seven countries decompose objectives hierarchically into sub-objectives. We tested it in a few places the hypothesis that large nations employ more fully developed decision frames, producing more agreement with framing activity statements than smaller nations. For the question on developing preferences into criteria to evaluate PIP merit, a Mann-Whitney one-tailed test shows almost significant support ($p = 0.058$). Overall, while there appears to be appreciation for the utility of developing and applying detailed and comprehensive portfolio objectives in terms that can be measured, the responses and comments of nations reveal multiple challenges in fully implementing this approach.

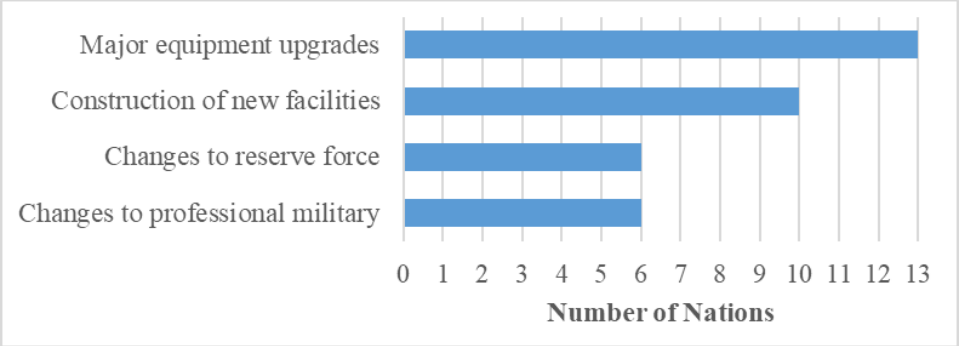

Figure 1: Other resource allocations prioritized with capital equipment investments in the same decision process.

Figure 2: National responses to DIP framing behaviour statements.

The survey probed thirteen different limited resources that could, in turn, limit PIP content. Most frequently Identified PIP constraints were investment funding (9 nations) and operating budget (5 nations). The most common other resource limitations were merely discussed: facilities (9), project management capacity (8) and training (8).

4.2 Feasible Alternatives

It is hard to satisfy some portfolio objectives when few projects make any contribution. Recognizing and correcting this with new investment options significantly increases decision maker freedom to prioritize [10]. Figure 3 reports the prevalence of this practice among nations surveyed.

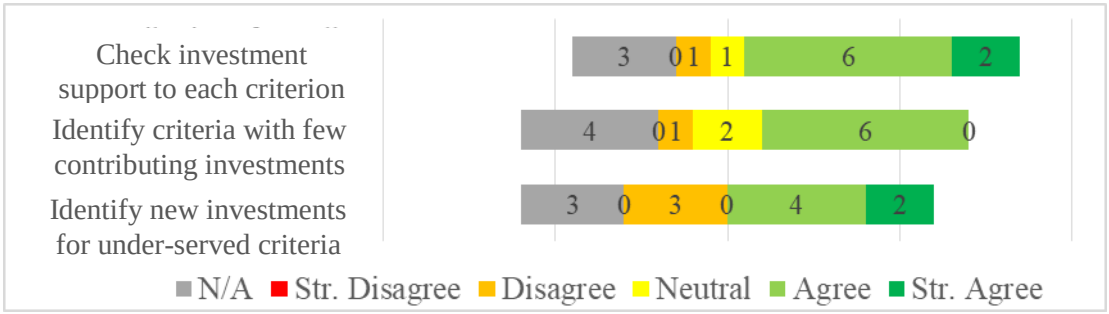

Figure 3: Responses to statements on available investment support for portfolio benefit criteria.

As shown in Figure 3, eight nations agreed that they examine how various candidate investments contribute to their benefit criteria, and three indicated their country did not explicitly use PIP benefit criteria. Interestingly, only six respondents indicate they recognize potential benefit gaps that might exist in a PIP when too few candidate investments exist to help fill a specific benefit criterion. Six respondents also indicate they actively search for new candidate investments to fill those under-served PIP benefit criterion. Accompanying comments by several countries describe consideration of capabilities and gaps without explicitly mapping a portfolio to a well-defined set of benefit criteria.

4.3 Values and Trade-offs

Values and trade-offs enable overall evaluations of alternative PIPs. Survey questions were designed to reveal whether nations tend to use a more quantitative or qualitative approach.

While Figure 4 indicates seven nations agreed they develop and use metrics reflecting portfolio satisfaction of PIP criteria, accompanying comments suggest some nations only aspire to use quantitative metrics or that the only metric is cost.

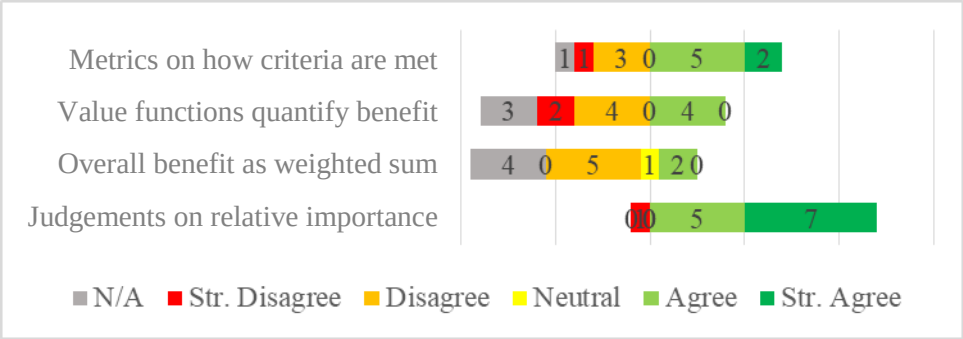

Figure 4: Approaches for evaluation of alternative investment portfolios.

Responses to subsequent questions in Figure 4 suggest most nations capture alternative portfolio benefits qualitatively. Nearly all (12) nations elicit from decision makers and participating stakeholders judgments regarding the relative importance of different criteria. While four nations report they combine metrics using a value function, and two that use a weighted sum to measure benefits, comments suggest these approaches are used for individual projects, but not the entire portfolio. We hypothesized that larger nations have more numerous investments, larger investment portfolio budgets and greater decision support, and therefore employed more quantitative benefit modelling than smaller nations, which would be reflected in more agreement with statements of quantitative value development. A Mann-Whitney U test of responses to related statements support the hypothesis at a significance level of 0.048, although comments that accompanied some agreeing responses about aspiration and use of only cost metrics undermine this marginal result.

The survey presented respondents with sixteen tools and techniques from the literature to determine the extent to which they are used to model the benefit expected from a particular portfolio. More than half the countries (8 out of 12)[5] report at least one tool or technique as being either “directly related” or “used” by that country in its benefit modelling. The following four tools were the only ones that are “used” or “directly related” by more than a single country

- Priority lists (1 used, 3 directly related)
- Multi-Objective / Multi-Criteria Decision Analysis / Decision Making (0 used, 4 directly related)
- Requirements management (2 used, 0 directly related)
- Additive value (1 used, 1 directly related)

Three nations were unfamiliar with more than one of the techniques and two were unfamiliar with 6 or more of them. Based on rather high number of responses indicating unfamiliarity, it appears tools and techniques drawn from the literature for this section of the survey are not well known among a few survey respondents fulfilling the selection criteria. Examining the results and accompanying comments suggest that quantitative tools and techniques are more used in project than portfolio management. The most unfamiliar items were Outranking methods (5) and Swing weights (3).

4.4 Information

The survey asked which of twelve cost categories are included in PIP budgets. The cost categories, in descending order of responses, were purchase cost, supporting IT systems, investment R&D, transport of deliverables and provisions, project management, facilities, mid-life upgrade and annual operating costs, operator training, transition costs (facilities conversion, new tactics & training development), end-of-life

[5] One nation did not complete this section.

costs and operating personnel costs. Cost estimates were roughly evenly split between point estimates and range estimates that incorporate uncertainty.

All twelve respondents[5] include the purchase of new systems in their PIPs. For those two nations that did not include R&D costs, the accompanying comments suggest some equipment is purchased "off-the-shelf" so that R&D costs may already be incorporated in the purchase price. Surprisingly, only seven respondents indicate they include future operating costs in their PIP budget estimates. Since PIP time frames are generally shorter than the expected operational life of new systems, it is not surprising only four nations factor disposal (End-of-Life) costs/benefits in their PIP estimates.

Risk assessment models (e.g. [11][12]) in the literature typically address four questions: 1) What can go wrong? 2) What are the chances? 3) What are the consequences and 4) What needs to be done? In evaluating individual candidate investments in a PIP, there are three main concerns: a candidate investment may cost more than estimated (cost); a candidate investment might be available later than expected (schedule); and a candidate investment may deliver less value than anticipated (performance). Figure 5 shows survey responses to statements that these risks are modelled for candidate investments as well as the risk of portfolio budgets being lower than expected.

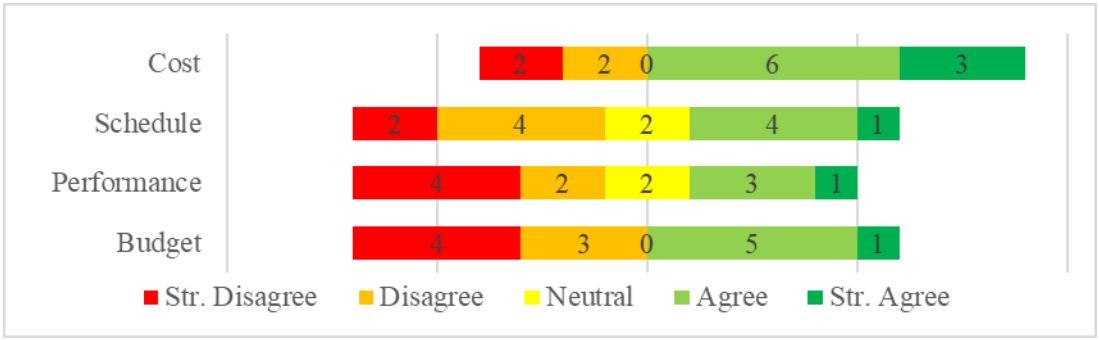

Figure 5: Responses to statements that investment and portfolio budget risks are modelled.

Costs are modelled by nations more often than other risks. In the accompanying comments many nations recognized and discussed a variety of risks in their PIP decision process, but indicated that they mostly do not explicitly model those risks, nor assign specific probabilities to assess their likelihood. Comments among nations that do not model a specific type of risk, indicate that they still have contingency planning.

4.5 Sound Reasoning

The survey explored reasoning in PIP reviews by asking about use of portfolio optimization software and looking into various kinds of interactions between candidate investments. More than half of respondents (from 4 larger and 4 smaller nations) indicated they use optimization. Possible interactions to model include investments that are: complements i.e. one investment requires another (or others) to operate successfully. Another interaction is between substitutes i.e. one investment can be used to fulfill the same requirements as another. Other possible interactions exist where investments offer: Positive externalities i.e. one investment offers spillover benefits that satisfy other requirements; or synergies i.e. the benefit from combined investments is greater than the sum of their individual benefits. Another possible interaction is where two or more investments linked by dependency where the success of one is linked to the success of another.

The most often modelled interaction type is substitute investments. It is formally modelled by five nations as constraints disallowing more than one, followed by complements (four nations) either as constraints to include or value penalties when complements are not selected. The four other kinds of interactions are formally modelled by only few nations. Instead of explicitly modelling interactions, most nations indicated

that they deal with each kind of possible candidate investment interactions in Figure 6 by noting the interaction in discussions and adjusting the planned investment portfolio accordingly. However, there is a small and somewhat variable number of nations that ignore some forms of interactions.

There was no clear difference in treatment of alternative investment interactions based by nation size. For example, of the four nations ignoring linked execution, three were small and one large, not really notable. While respondents generally consider investment interactions to be important in the PIP process, they tend not to use formal models or constraints to help address those interactions. There appears to be consensus that, while investment interactions are important, it is difficult to capture all the complexities and are mostly addressed through discussion, perhaps supported by prior capability based planning models and other analyses.

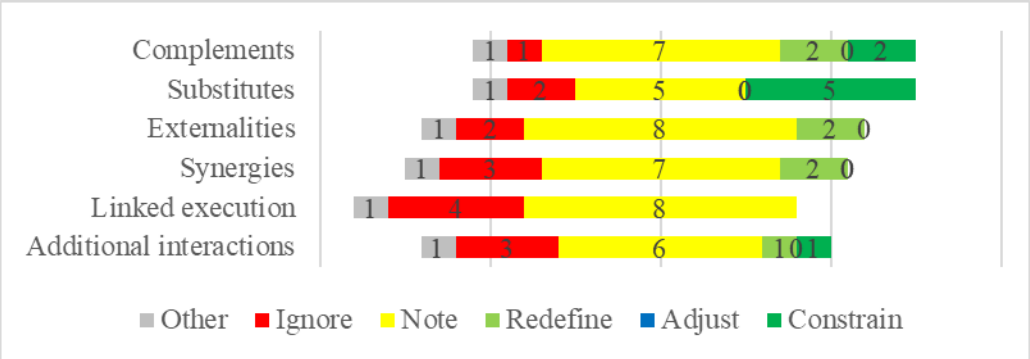

Figure 6: Responses preferred to value and risk interactions recognized between investments.

4.6 Commitment to Act

The survey asked respondents to outline the stages of their PIP decision processes. While each response was unique, in the following a distillation of several common PIP development steps is described below.

The first stage typically involves planning guidance instructions from the Ministry of Defence. Each authority or component establishes their requirements with some awareness of constraints. Assessments are conducted on what has changed since the last review and data is collected on investment alternatives. The second stage often links Plans and Desired End-States. Military departments submit proposals for additions and modifications along with system requirements. A second round of discussions matches requirements and financial resources (trade-offs between capabilities, in terms of quantity and quality), along with Strategic Reviews of External Risks. The draft PIP including materiel acquisitions and expected capabilities is based on operational requests and rough cost estimates. A third stage involves a capability, concepts and program review by senior leadership and analysis of operating budget requirements. Initial balancing of the PIP, and base lining it to the budget, is followed by discussions of portfolio options with leadership and revisions as required. A fourth stage prioritizes budget allocations. A draft PIP may assign priorities and establishes funded programs, and those in reserve. In some cases, a government committee endorses the PIP and sends it to inter-ministerial authorities (ministry of finance, etc.), possibly for comment by the executive branch, and the treasury. The later PIP stages can involve the final development of portfolio options, careful cost estimation for the first budget year, and a detailed budget review. Final reviews including military service chiefs, and technical and legislative Departments (where new requirements may emerge and require PIP re-prioritization) are conducted leading to approval by the Minister of Defence.

The survey also asked about the approaches used by nations for generating shared decision commitment and stakeholder involvement. In particular, the response menu offered Dialog Decision Process and Decision

Conferencing, possibly with Consensus Modelling. Ten nations specifically reported using a Dialog Decision Process, and four report using Decision Conferencing. Two nations use Consensus Modelling. Some reported having used more than one process. Given a sufficiently broad interpretation of Dialog Decision Process (to include aspects of Decision Conferencing and Consensus Modelling) it appears most group decision processes might be said to follow a generalized Dialog Decision Process approach. This suggests these nations have a sequence of meetings involving teams made up of decision-makers, stakeholders and analysts, who together are involved in PIP development. The survey also suggests significant differences in process maturity, including the use of formal methods, how well roles and various steps in the process are defined, and the extent to which stakeholder preference data is recorded for reference.

5.0 CONCLUSIONS

The results of the survey show that every responding nation did have a planned investment portfolio. This portfolio included major capital investment, and was constrained by available resources, most often by available funding. All responding nations could explicate a process for drafting and approving PIP while the process itself was unique in each nation. With a sufficiently broad interpretation of Dialogue Decision process, most group decisions can be said to follow a generalised Dialog Decision Process approach.

All but one nation elicit judgments regarding the relative importance of different criteria from decision makers and participating stakeholders. Ten countries solicit preferences and priorities from PIP decision makers. Comments by several countries describe consideration of capabilities and gaps without explicitly mapping a portfolio to a well-defined set of portfolio benefit criteria. Most nations capture alternative portfolio benefits qualitatively but larger nations appear to employ more quantitative benefit modelling.

The countries were divided with respect to developing a hierarchical decomposition of objectives to sub-objectives. In terms of interaction modelling, a minority of countries formally modelled substitute and complementary investment candidates while other kinds of interactions between investments were modelled only by a few nations. Instead, most nations indicated that they deal with each kind of possible candidate investment interaction by noting the interaction in discussions and adjusting the planned investment portfolio accordingly. The countries were also divided with respect to including force size decisions in their PIP process corresponding to “Total System Cost”.

Of the sixteen benefit modelling tools & techniques surveyed, only priority lists, MCDA, Requirements Management and Additive Value were either used or directly related to current practice by more than one nation. However, more than half of nations at least one technique was used or at least directly related. While countries were also divided in their responses to modelling investment risks, in the comments many nations indicated that they mostly do not explicitly model those risks, nor assign specific probabilities to assess their likelihood but they discuss a variety of risks in their PIP decision process.

The survey results and accompanying comments also suggest that portfolio benefit-maximizing strategies that are employed tend to be more ad hoc and focus on deliberation without modelling and formal algorithmic/mathematical support. A related trend seen throughout the survey is that response distributions, both shown here and those excluded from the paper for brevity, show responses largely agreeing with statements about initial steps progressing through a technique but shifting toward disagreement with statements about later stages. This resonates with how few named benefit modelling techniques were employed by nations in their portfolio analyses. Clearly, computational techniques are in abundant supply in the literature, but show very limited uptake to support real PIP decisions, just as the problems formulated in the literature do not acknowledge the political complexities or real investment prioritization. The survey confirms the sense of mismatch between technique offered by the preponderance of investment prioritization literature and national uptake of those offerings.

Beyond this trend, the dominant observation from almost every part of the survey is how varied national practices appear to be. Procedures vary widely between nations for natural enough reasons, having been shaped by distinct cultures, histories and philosophies of government. The most significant insight from this has been the likely futility of pointing to one nation's approach as exemplifying superior practice. This will seem (and be) somewhat arbitrary and out of context to other nations.

These observations confirmed the emerging recognition of a need for investment prioritization guidance to be principles-based rather than recommending procedures. Defence investment prioritisation is among the set of most complex recurring national problems, only somewhat due to computational complexity and principally because of the problem context and organizational complexity (more human complexities). It was in this light that the utility of the Decision Quality construct became most apparent, motivating its specific application to the investment selection context, which has proven to be the core of our work.

As a final note, we expect there to have been some inevitable misunderstanding of question meanings to an extent difficult to estimate. Even the term "national practice" may exaggerate the actual degree of regularity of national approach and conceal gradual iteration between the infrequent and intrinsically complex process instances. Survey result significance will be better understood after a follow-on research workshop being planned for 2023 intended for nations to present their processes in their own unclassified terms.

6.0 REFERENCES

[1] Heidenberger, Kurt and Stummer, Christian (1999). "Research and development project selection and resource allocation: a review of quantitative modelling approaches." *International Journal of Management Review*s 1(2): 197

[2] Brown, Gerald G., Dell, Robert F. and Newman, Alexandra M. (2004). "Optimizing Military Capital Planning." Interfaces 34(6): 415-425

[3] Harrison, Kyle Robert, Elsayed, Saber, Garanovich, Ivan, Weir, Terence, Galister, Michael, Boswell, Sharon, Taylor, Richard and Sarker, Ruhul (2020). "Portfolio Optimization for Defence Applications." IEEE Access 8,1 (2020): pp. 60152-60178

[4] Brown, Gerald G and Rosenthal, Richard E. (2008). "Optimization Tradecraft: Hard-Won Insights from Real-World Decision Support." Interfaces 38(5): 356-366.

[5] Howard, Ron (1968). "Decision Analysis: Applied Decision Theory" in *Proceedings of the Fourth International Conference on Operatinoal Research*. New York: John Wiley, pp. 55-71.

[6] Spetzler, Carl, Winter, Hannah, and Meyer, Jennifer (2016). *Decision Quality: Value Creation from Better Business Decisions*. Hoboken, NJ: John Wiley & Sons, Inc.

[7] Lykke Jr., Arthur F. (1989). "Defining Military Strategy" in Military Review 69:5 (May 1989).

[8] Tate, David M. and Thompson, Paul M. (2016), "Portfolio Selection and Resource Allocation for Defence Applications," Institute for Defence Analyses, IDA Document NS D-5439, March 2016.

[9] Tate, David M. and Thompson, Paul M. (2017), "Portfolio Selection Challenges in Defence Applications," Institute for Defence Analyses, IDA Document NS D-8493, August 2017.

[10] Keeney, Ralph (1992). *Value-Focused Thinking: a Path to Creative Decisionmaking*. Cambridge MA: Harvard University Press.

[11] ISO 31000 (2018) Risk Management -- Guidelines.

[12] NATO Research Task Group SAS-109, Risk Assessment Guidebook for Defence Acquisition Programmes, STO Technical Report SAS-109, 2018.

ACKNOWLEDGEMENTS

The authors would like to thank the other members of the SAS-134 team for their contribution to SAS-134 effort that forms the basis for this paper. In alphabetical order: Alf-Christian Hennum, Francois Melese, Anthony Ween and Petráš Zdeněk, and past member Anke Richter.

BIOGRAPHIES

John Steele earned his PhD from the University of Saskatchewan in 2002 for work in numerical analysis and convergence acceleration and has worked since then as an operational research scientist with Defence Research & Development Canada's Centre for Operational Research & Analysis. He has served in postings supporting Canadian Army experimentation in Kingston, Royal Canadian Air Force capability management in Winnipeg and, since 2012, strategic defence planning and management in Ottawa. His avocations include operatic singing (having previously earned a Bachelor of Music in voice performance) and recreational soccer.

Juha-Matti Lehtonen earned his doctoral degree in 1999 at Helsinki University of Technology (HUT) for work on supply chain modelling. During the 2000's he worked in industry twice, at Delfoi Oy and Winmodal Oy and also as a teaching researcher at HUT. In 2007 he became a professor of operations management & strategy at Tampere University of Technology. In 2011 he moved to his current employment as a professor of Defence Acquisition at the National Defence University. His current research interests involve military procurement and acquisition. His personal interest revolve around family, marathon and a keen interest in history.